JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Possible observation of electromagnetic cascades in extragalactic space

A. V. Uryson

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

(Submitted 1 November 1996)
Zh. Éksp. Teor. Fiz. **113**, 385–397 (February 1998)

Arrival directions of gamma-ray-initiated showers with energies over 10^{14} eV detected by the Bolivian and Tien Shan high-altitude arrays have been analyzed. Their distribution over the celestial sphere is nonuniform, and in the range of galactic latitudes $b \geqslant 30°$ it is similar to the distribution of Seyfert galaxies, which are at distances ~1.5–200 Mpc from us, if the Hubble constant is 75 km/s·Mpc. Assuming that Seyfert galaxies are sources of protons with energies higher than 3×10^{19} eV, the gamma-rays can be generated in collisions of extragalactic protons with relict photons and in subsequent electromagnetic cascades in the extragalactic space. The upper limit on the extragalactic magnetic field, $B \ll 10^{-9}$ G, is derived.
[S1063-7761(98)00102-4]

1. INTRODUCTION

It was predicted[1,2] that interaction between protons and relict radiation in the extragalactic space can manifest itself through a cutoff in the spectrum of cosmic rays in the energy range $E > 3\times10^{19}$ eV. But the spectrum should not have a black-body cutoff if sources of cosmic protons are located at relatively small distances, within several tens of megaparsecs, from us.[3] As shown in Ref. 4, it is possible that protons with energies $E > 3\times10^{19}$ eV are generated in nuclei of active galaxies located at distances within 40 Mpc at the Hubble constant of 75 km/s·Mpc. If this is the case, the proton spectrum should not have a black-body cutoff.

The present study shows that gamma-rays of ultrahigh energies detected by ground-based arrays can also be a manifestation of interaction between protons and relict radiation.

Gamma-ray-initiated showers were investigated on the Tien Shan[5,6] and Bolivian[7,8] high-altitude arrays. Against the background of numerous showers initiated by cosmic rays, these showers could possibly be distinguished by their poor muon and hadron content. The researchers working on the Bolivian array identified them using this feature. On the Tien Shan array, the basic criterion for selection of showers was the low energy flux carried by the hadron component as compared with electrons and photons.

The energies of gamma-quanta detected by the Tien Shan array[5,6] were $(5-8.5)\times10^{14}$ eV, on the Bolivian array[7,8] the energies were 10^{14}–10^{17} eV. The arrival directions of these showers did not point at any galactic sources of γ-radiation. The measured fluxes were several dozen times higher than the theoretical estimates of diffuse γ-radiation generated in interaction between cosmic rays and extrastellar gas.[9,10] For this reason, the detected gamma-rays were interpreted as radiation from unresolved galactic sources.[5–8,11,12] It follows that large numbers of point sources should be concentrated in small sections of the celestial sphere. In this case they should have been detected in the x-ray band, but no excessive x-ray emission from the areas with an allegedly higher density of sources have been observed.[10] Gamma-ray initiated showers have not been investigated.

This paper gives an analysis of the distribution of arrival directions of gamma-ray-initiated showers over the sky and suggests a model for interpretation of features of this distribution and intensity of this radiation.

2. DISTRIBUTION OF GAMMA-RAY INITIATED SHOWERS OVER THE CELESTIAL SPHERE

Let us consider the distribution of shower arrival directions over the celestial sphere in the coordinates (α,δ) given in Ref. 5. It is shown in Fig. 1. The numerical labels attached to the showers were suggested by the author. The distribution has some peculiar features. First, the coordinates of gamma-rays labeled by 2, 3, 4, and 6 coincide within their 1.5-fold measurement errors. The arrival direction of the gamma-ray labeled by 5 is within the three-fold error of these four showers. As follows from the statistics, the actual arrival direction of a gamma-ray photon is within the three-fold measurement error with a probability of 99.8%. The probability of an accidental coincidence of five gamma-ray photons among eight detected events can be estimated in the following ways. These gamma-rays arrived from the ascension band $\alpha = (13-16)$ h. The probability P that five gamma-ray photons among eight fall within a band of width $\Delta\alpha = 3$ h is rather small:

$$P = C_8^5(\Delta\alpha/24)^5(1-\Delta\alpha/24)^3 \approx 1.1\times10^{-3}.$$

Moreover, these gamma-rays arrived from an area that occupies a fraction $\Delta S \approx 0.19$ of the total field of survey. The probability that five out of eight gamma-ray photons fall within this area accidentally is also quite small:

$$P = C_8^5(\Delta S)^5(1-\Delta S)^3 \approx 7\times10^{-3}.$$

Second, gamma-rays Nos. 2–6 arrived from the range of galactic latitudes $b \geqslant 30°$. This means that they are probably not diffusive, since the intensity of diffuse emission in these directions is lower than along the galactic disk,[10] and most

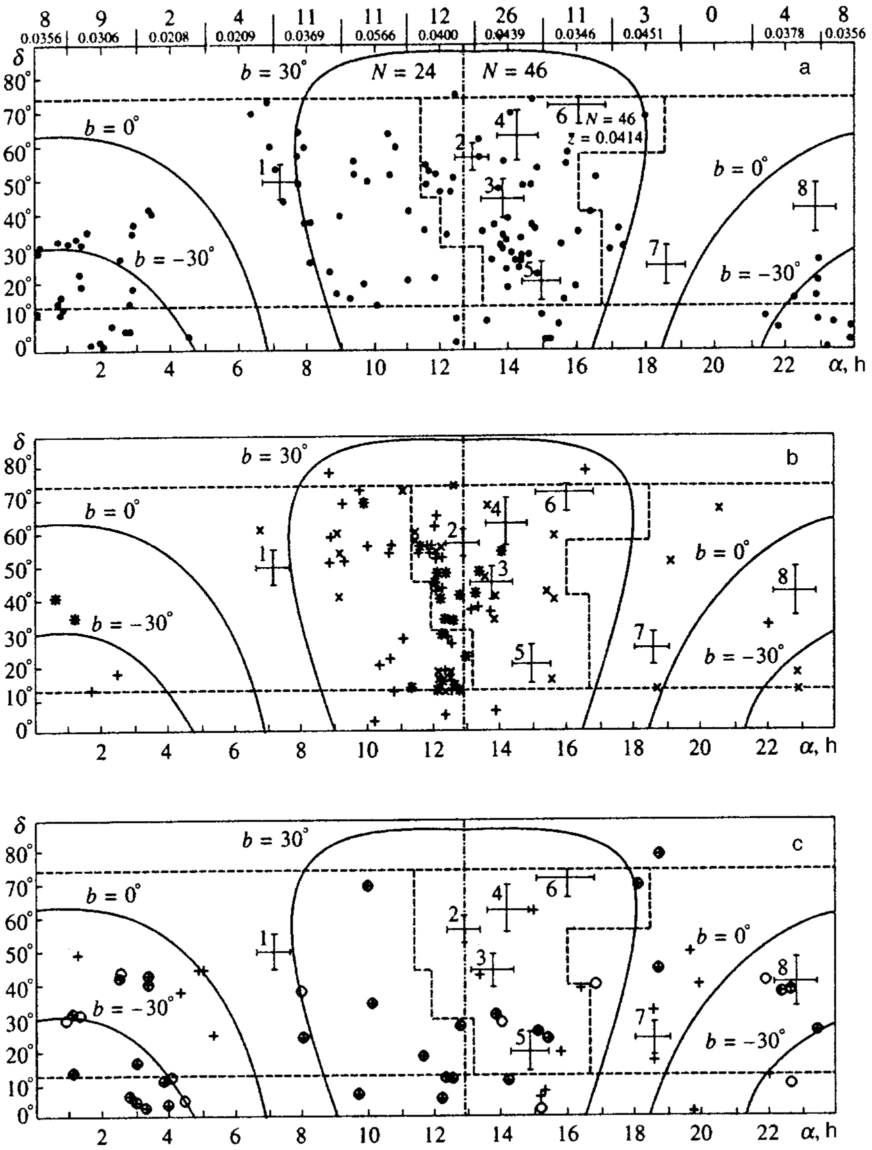


FIG. 1. Distribution of arrival directions of showers due to gamma-ray photons on the sky.[5] The solid lines show galactic latitudes $b=\pm 30°$ and the galactic plane $b=0°$. The dashed lines limit the observation area of the Tien Shan array and the area of photons Nos. 2–6 plus the three-fold measurement errors in their coordinates. The dash-dotted line $\alpha=12$ h 49 min is the axis of symmetry of the region $b>30°$. (a) Distribution of Seyfert galaxies from the full Byurakan survey.[18] Numbers N of galaxies and their $\overline{z}$ in the band $\Delta\alpha=2$ h and in the region (2–6) are given, as well as numbers N in the region $b>30°$ on the right and on the left of line $\alpha=12$ h 49 min; (b) distribution of Seyfert galaxies and active galactic nuclei with active cores listed in Refs. 19 and 20 with $z\leqslant 0.0022$ (*), $0.0022<z\leqslant 0.0049$ (+), and $0.0049<z\leqslant 0.0089$ (×); (c) distribution of powerful radio galaxies with $z<0.1$ at frequencies of 178 MHz[21] (+) and 5 GHz[22] (○).

gamma-ray photons should have been expected from the region of lower latitudes. (Absorption of gamma-rays of these energies in the galactic disk can be neglected.[10]) In other experiments[8] the excess (on the level of 3.8σ) of gamma-ray-initiated showers with energies higher than 10^{14} eV was detected in the region $\alpha=180-210°$, $\delta=0-(-40°)$. This region also corresponds to high latitudes $b<-25°$ (the distribution of celestial coordinates of these showers was not published). The arrival directions of showers in regions of high galactic latitudes indicate that these gamma-rays can be of extragalactic origin: they could be generated in interactions between extragalactic protons with the relict radiation and in subsequent electromagnetic cascades in the extragalactic space. This assumption was formulated earlier.[13]

3. DISTRIBUTION OF POSSIBLE SOURCES OF COSMIC PROTONS OF SUPERHIGH ENERGIES

Feasible sources of high-energy protons are probably nuclei of active galaxies[4] and/or powerful radiogalaxies.[10,14] Let us suppose that protons with energies $E>3\times 10^{19}$ eV propagate in extragalactic space along straight lines and let us compare the distribution of their sources over the celestial sphere with the distribution of gamma-rays.[5] The uncertainty in optical coordinates of galaxies is within several angular seconds, so the errors in the distribution of galaxies over the sky are minor.

The analysis of the distribution of Seyfert galaxies was based on the full Byurakan survey.[18] The number of galaxies in this survey was $N_{\mathrm{FBS}}=127$ and the average red shift was $\overline{z}_{\mathrm{FBS}}=0.048$. Their distribution is given in Fig. 1a. In order to examine possible nonuniformity in the distribution, the region $b>30°$ was divided by the line $\alpha=12$ h 49 min into two sections of equal areas and the number of galaxies in each section was counted. The line $\alpha=12$ h 49 min is the symmetry axis for the region $b>30°$ in the coordinates (α,δ). Moreover, galaxies were counted in the bands $\Delta\alpha=2$ h and in the region of arrival directions of gamma-rays Nos. 2–6 within the three-fold measurement error. Numbers of this galaxies are given in Fig. 1a. It follows from the figure that the distribution of galaxies in region $b>30°$ is nonuniform, and the region of showers Nos. 2–6 coincides with that where the number of galaxies is largest.

TABLE I. Numbers of galaxies with $z \leqslant 0.0089$ in different sections of the sky.[19,20]

	Number N of galaxies			
	$z \leqslant 0.0089$	$z \leqslant 0.0022$	$0.0022 < z \leqslant 0.0049$	$0.0049 < z \leqslant 0.0089$
(1) $\alpha = (0-24)h$, $\delta = (0-90°)$	67	24	32	26
(2) Region $b > 30°$ in the field of survey of the array	59	20	23	16
(3) Region (2–6)	38	13	14	11
Fraction of galaxies in (3) relative to their number in (1)	0.57	0.54	0.44	0.48
Fraction of galaxies in (3) relative to their number in (2)	0.64	0.65	0.61	0.68

Note. Region (3) has an area of about 0.19 of area (1) and about 0.52 of area (2).

The distribution of Seyfert galaxies with $z \leqslant 0.0089$ from the catalogue[19] which could also be used in statistical studies was analyzed in addition. The distribution also included active galactic nuclei from Ref. 20 not included in Ref. 19 (the galaxies from Ref. 20 did not fall with the region of showers Nos. 2–6). The distribution of galaxies is shown in Fig. 1b. Galaxies with different z were counted on different sections of the sky: (1) $\alpha = 0-24$ h, $\delta = 0-90°$; (2) the band $b > 30°$ in the field of survey of the Tien Shan array; (3) the region of showers Nos. 2–6. Their numbers and the fraction of galaxies in region 2–6 relative to the numbers of galaxies in area (1) and (2) are given in Table I. The table also indicates that a notable fraction of near galaxies falls within the region 2–6.

In the analysis of the distribution of powerful radio galaxies over the sky, the catalogues for frequencies of 178 MHz and 5 GHz[21,22] were used. It turned out that the number of galaxies with $z < 0.1$ in the region $b > 30°$ to the right and to the left of the line $\alpha = 12$ h 49 min is different: $N = 14$ and $N = 8$. In the region of showers Nos. 2–6, however, their number equals the number of galaxies in the areas $b > 30°$ free of gamma-rays. The distribution of radiogalaxies with $z < 0.1$ is given in Fig. 1c.

It follows that, if gamma-rays Nos. 2–6 are of extragalactic origin, the main sources of protons generating them are, apparently, Seyfert galaxies rather than radio galaxies. This conclusion is in agreement with Ref. 4.

4. POSSIBLE MECHANISM OF GAMMA-QUANTA GENERATION

Let us consider the development of electromagnetic cascades[15–17] and distances from proton sources at which gamma-rays with energies of $(5-8.5) \times 10^{14}$ eV are generated in them. The mean free path of such gamma-rays in the field of relict photons is 10–20 kpc,[23] so they should be generated on the boundaries of the Galaxy in order to be detected on the Earth. Numerical estimates given in the paper refer to electromagnetic cascades generated by protons with energies $E = 6 \times 10^{19}$, 3×10^{20}, and 10^{22} eV.

Protons with energies $E > 3 \times 10^{19}$ eV interact with relict photons through the following reactions:

$$p + \gamma_{\mathrm{rel}} \rightarrow p + \pi^0, \tag{I}$$

$$p + \gamma_{\mathrm{rel}} \rightarrow n + \pi^+. \tag{II}$$

In the proton rest system, the cross section $\sigma_{p\gamma}$ has a resonance at the photon energy $\varepsilon' \approx 300$ MeV with the peak $\sigma_{p\gamma}(300\ \mathrm{MeV}) \approx 5.5 \times 10^{-28}\ \mathrm{cm}^2$ and drops with ε' to $\sigma_{p\gamma}(5\ \mathrm{GeV}) \approx 0.8 \times 10^{-28}\ \mathrm{cm}^2$.[24] The photon energy is $\varepsilon' = \gamma\varepsilon(1 - \beta \cos \vartheta)$, the proton Lorentz factor is $\gamma = E/m_p$, the proton mass is $m_p \approx 938$ MeV, the energy of the relict photon in the L-system is $\varepsilon \approx 6.7 \times 10^{-4}$ eV, $\beta = \sqrt{1 - 1/\gamma^2}$, and ϑ is the angle between the proton and photon momenta in the L-system.[3] Since we have $\beta \approx 1$ and $-1 = \leqslant \cos \vartheta \leqslant 1$, then $0 < \varepsilon' < 2\gamma\varepsilon$. In a head-on collision ($\cos \vartheta = -1$) the value $\varepsilon' = 300$ MeV corresponds to the energy $E \approx 3 \times 10^{20}$ eV and $\varepsilon' = 5$ GeV to $E \approx 10^{22}$ eV. The proton mean free path is $\bar{\lambda}_{p\gamma} = (n_{\mathrm{rel}} \sigma_{p\gamma})^{-1}$, where the density of relict photons is $n_{\mathrm{rel}} \approx 400\ \mathrm{cm}^{-3}$.[10] The mean free path $\bar{\lambda}_{p\gamma}$ can vary in the range $1.5 \leqslant \bar{\lambda}_{p\gamma} < 150$ Mpc for $E = 3 \times 10^{20}$ eV and $10 \leqslant \bar{\lambda}_{p\gamma} < 150$ Mpc for $E = 6 \times 10^{19}$ eV and $E \geqslant 3 \times 10^{20}$ eV. The fraction K_p of energy transmitted from the proton to the pion in reactions (I) and (II) increases with ε': $0.126 \leqslant K_p \leqslant 0.4$ at $\varepsilon' = 145$ MeV–2 GeV.[3] Let us put in the estimate the mean value $\bar{K}_p = 0.2$ for $E \ll 10^{22}$ eV and $\bar{K}_p = 0.4$ if $E = 10^{22}$ eV.

Mesons generated in reactions (I) and (II) decay:

$$\pi^0 \rightarrow 2\gamma, \tag{III}$$

$$\pi^+ \rightarrow \mu^+ + \nu, \tag{IV}$$

$$\mu^+ \rightarrow e^+ + \nu + \nu. \tag{V}$$

It follows from the kinematics of these decays[25] that the mean gamma-ray photon energy in Eq. (III) is $\bar{\varepsilon}_\gamma = \varepsilon_\pi/2$,

TABLE II. Energies of gamma-rays, $\overline{\varepsilon}_\gamma$, generated in reactions (I), (III), and limiting values of energies ε_1 and ε_2 of positrons, generated in reactions (II), (IV), (V), for several energies of the primary proton.

E, eV	$\overline{\varepsilon}_\gamma$, eV	ε_1, eV	ε_2, eV
6×10^{19}	6×10^{18}	1.5×10^{14}	1.2×10^{19}
3×10^{20}	3×10^{19}	8×10^{14}	6×10^{19}
10^{22}	2×10^{21}	5×10^{16}	4×10^{21}

TABLE III. Estimates of mean free paths of positron, $\overline{\lambda}_{\mathrm{IC}}$, in fields of relict and radio photons.

	$\overline{\lambda}_{\mathrm{IC}}$		
ε_e, eV	Field of relict photons	Field r_2	Field r_1
1.5×10^{14}	5 kpc	5 Mpc	17 kpc
8×10^{14}	5 kpc	5 Mpc	17 kpc
5×10^{16}	70 kpc	5 Mpc	17 kpc
1.2×10^{19}	6 Mpc	122 Mpc	17 kpc
6×10^{19}	40 Mpc	510 Mpc	17 kpc
4×10^{21}	2100 Mpc	2.1×10^4 Mpc	470 kpc

where the pion energy $\varepsilon_\pi = K_p E$, the muon energy in Eq. (IV) is in the range $(m_\mu/m_\pi)^2\varepsilon_\pi \leqslant \varepsilon_\mu < \varepsilon_\pi$, the positron e^+ energy in Eq. (V) is in the range $(m_e/m_\mu)^2\varepsilon_\mu \leqslant \varepsilon_e < \varepsilon_\mu$ or $(m_e/m_\pi)^2 K_p E \leqslant \varepsilon_e < K_p E$; the muon mass is $m_\mu \approx 106$ MeV, the charged pion mass is $m_\pi \approx 140$ MeV, and the electron mass is $m_e \approx 0.5$ MeV.

The values of $\overline{\varepsilon}_\gamma$ and of limiting positron energies, $\varepsilon_1 \leqslant \varepsilon < \varepsilon_2$ derived from kinematic relations are listed in Table II.

The decay path of pions and muons is less than 1000 m, so we assume that they decay instantaneously.

First let us follow the positron (hereinafter we follow only such particles that can generate gamma-rays with energies $\varepsilon_\gamma \geqslant 5\times10^{14}$ eV). It scatters from relict photons, and from radio photons if $\varepsilon_e > 10^{15}$ eV holds:

$$e + \gamma_b \rightarrow e' + \gamma'. \tag{VI}$$

The cross section of reaction (VI) is

$$\sigma_{\mathrm{IC}} = (3/8)\sigma_T q \,\ln(2/q + 0.5), \quad \text{if} \quad \varepsilon_e > \varepsilon_{\min} \approx m_e^2/\varepsilon_b;$$

$q = m_e^2/\varepsilon_e\varepsilon_b$, $\sigma_T = 6.65\times10^{-25}$ cm^2 is the Thomson cross section, and ε_b is the background photon energy. The energy of radio photons is $\varepsilon_b = 3\times10^{-9}-10^{-6}$ eV, and the energy density is $w_b \approx 10^{-7}$ eV/cm^3.[17] In our estimates, we consider a field r_1 of radio photons with $\varepsilon_b = 3\times10^{-9}$ eV, $n \approx 30$ cm^{-3} and field r_2 with $\varepsilon_b = 10^{-6}$ eV, $n \approx 0.1$ cm^{-3}. For the relict radiation field and radio photon fields r_1 and r_2, we have $\varepsilon_{\min} \approx 4\times10^{14}$, 8.3×10^{19}, and 2.6×10^{17} eV, respectively. The secondary positron energy in Eq. (VI) is $\varepsilon_{e'} \sim \varepsilon_{\min}$ and the scattered gamma-ray photon energy is $\varepsilon_{\gamma'} \sim \varepsilon_e$. If the proton energy satisfies $E \leqslant 10^{22}$ eV, the positron energy is in the interval

$$1.5\times10^{14}-10^{15} < \varepsilon_e < 6\times10^{18}-10^{20} \text{ eV}.$$

In the field of relict photons, a positron with energy $\varepsilon_e = 5\times10^{14}-10^{15}$ eV scatters a gamma-ray photon with an energy of $5\times10^{14}-10^{15}$ eV. The positron mean free path in the field of relict radiation is $\overline{\lambda}_{\mathrm{IC}} = (n_b\sigma_{\mathrm{IC}})^{-1}$ and equals $\overline{\lambda}_{\mathrm{IC}} \approx 5$ kpc for $\varepsilon_e = 5\times10^{14}-10^{15}$ eV. Thus, gamma-rays with energies of $5\times10^{14}-10^{15}$ eV should be generated in the field of relict photons in the single reaction (II) or in a single chain of reactions (IV), (V), (VI) if the proton energy satisfies $E \leqslant 10^{22}$ eV. This should happen at an average distance from the source $\overline{R} = \overline{\lambda}_{p\gamma} + \overline{\lambda}_{\mathrm{IC}}$ and $1.5 \leqslant \overline{R} < 150$ Mpc.

Moreover, gamma-ray photons with energies of $5\times10^{14}-10^{15}$ eV are scattered by positrons with energy $\varepsilon_e \approx 2\times10^{17}$ eV in the field r_1 and by positrons with $\varepsilon_e \approx 10^{16}$ eV in the field r_2. Here $\varepsilon_e < \varepsilon_{\min}$ holds and the scattering cross section is $\sigma_{\mathrm{IC}} \approx \sigma_T$, the mean energy $\overline{\varepsilon}_\gamma$ of scattered gamma-quanta and the positron energy ε being related by the following equation:

$$\overline{\varepsilon}_\gamma = (4/3)\varepsilon_b(\varepsilon_e/m_e^2).$$

Positrons of such energies are generated by protons with energies $6\times10^{19} < E \leqslant 10^{22}$ eV. The positron mean free path is $\overline{\lambda}_{\mathrm{IC}} = (n_b\sigma_T)^{-1} \approx 16$ kpc and 5 Mpc in fields r_1 and r_2, respectively. Therefore gamma-rays with energies of $5\times10^{14}-10^{15}$ eV are also generated in both reaction (II) and reaction chain (IV)–(VI) in the field of radio photons at distances $1.5 < \overline{R} < 155$ Mpc from a source of protons with energies $6\times10^{19} < E \leqslant 10^{22}$ eV.

Table III shows estimates of mean free path $\overline{\lambda}_{\mathrm{IC}}$ of positrons with energies ε_e listed in Table II in the fields of background photons. Secondary positrons with energy $\varepsilon_{\min}$ generated in reaction (VI) scatter gamma-ray photons with energies of $5\times10^{14}-10^{15}$ eV, as was demonstrated above. It follows from the data in Table III that they are created at distance $\overline{R} = \overline{\lambda}_{p\gamma} + \overline{\lambda}_{\mathrm{IC}}$ from the source, and we have $1.5 < \overline{R} < 200$ Mpc for $E \leqslant 10^{22}$ eV and $\overline{R} \leqslant (1-2)\times10^3$ Mpc for $E = 10^{22}$ eV.

Now let us follow the gamma-rays generated in reactions (III) and (VI). In the field of background photons they generate electron–positron pairs if $\varepsilon_\gamma \geqslant \varepsilon_{\min}$:

$$\gamma + \gamma_b \rightarrow e^+ + e^-. \tag{VII}$$

The cross section defined by Eq. (VII) is

$$\sigma_{\gamma\gamma} = (3/8)\sigma_T a^2[(2 + 2a^2 - a^4)\ln(a^{-1} + \sqrt{a^{-1} - 1}) - \sqrt{1 - a^2}(1 + a^2)],$$

where $a = m_e/\varepsilon_\gamma^*$ ($\varepsilon_\gamma^* \geqslant m_e$). The energy of the photons in the frame where their total momentum is zero equals

$$\varepsilon_\gamma^* = [(\varepsilon_\gamma\varepsilon_b/2)(1 - \cos\psi)]^{1/2},$$

where ψ is the angle between the photon momenta in the L-system. The energies of particles in the pair generated in reaction (VII) are $\varepsilon_{e_1} \sim \varepsilon_\gamma$ and $\varepsilon_{e_2} \sim \varepsilon_{\min}$. The mean free path of gamma-rays in Eq. (VII) is $\overline{\lambda}_{\gamma\gamma} = (n_b\sigma_{\gamma\gamma})^{-1}$. Estimates of $\overline{\lambda}_{\gamma\gamma}$ for gamma-rays with energies ε_γ listed in Table II are given in Table IV.

Electrons generated with energy $\varepsilon_e \sim \varepsilon_{\min} = 4\times10^{14}-2.6\times10^{17}$ eV scatter gamma-ray photons with energies of $5\times10^{14}-10^{15}$ eV in reaction (VI) as described above. These

TABLE IV. Estimates of mean free paths of gamma-rays, $\overline{\lambda}_{\gamma\gamma}$, in fields of relict and radio photons ($\psi=90°$).

	$\overline{\lambda}_{\rm IC}$		
ε_γ, eV	Field of relict photons	Field r_2	Field r_1
6×10^{18}	3.3 Mpc	-	-
3×10^{19}	13.6 Mpc	211 Mpc	-
2×10^{21}	667 Mpc	7100 Mpc	212 kpc

gamma-rays are generated at an average distance $\overline{R}=\overline{\lambda}_{p\gamma}+\overline{\lambda}_{\gamma\gamma}+\overline{\lambda}_{\rm IC}$ from the proton source. If a pair is generated in reaction (VII) in the field of relict photons, gamma-ray photons are created at distances from the source $1.5<\overline{R}<160$ Mpc at $E=6\times10^{19}$ eV, $100\leqslant\overline{R}<250$ Mpc at $E=3\times10^{20}$ eV, and $\overline{R}\leqslant10^3$ Mpc at $E=10^{22}$ eV. For $E=10^{22}$ eV a pair can be also created in reaction (VII) in the field r_2, and gamma-rays will be generated at distance $4<\overline{R}<155$ Mpc from the source.

The resulting estimate of distance $\overline{R}\approx200$ Mpc coincides with the distances to Seyfert galaxies listed in the full survey[18] which fall within the region where gamma-rays Nos. 2–6 were detected, if the Hubble constant $H=75$ km/s·Mpc. It also follows from our estimate that in the reactions given above gamma-rays with energies $(5-8.5)\times10^{14}$ eV can be generated at distances of $(1-2)\times10^3$ Mpc from the source.

As a result of the multiple reactions (I) and (II), which occur for $E>3\times10^{19}$ eV, and of reactions (III)–(VII), electromagnetic cascades are generated. Like external air showers,[26] these cascades may develop over distances considerably shorter than the average length because of fluctuations in the parameters ϑ, K_p, $\lambda_{p\gamma}$, $\lambda_{\gamma\gamma}$, $\lambda_{\rm IC}$, ε_e, and ε_{e_2}: the initial proton undergoes several collisions at angle ϑ corresponding to the maximum cross section $\sigma_{p\gamma}$ with energy transfer $K_p<\overline{K_p}$, the mean free path being $\lambda<\overline{\lambda}$, in reaction (V) $\varepsilon_e\ll\varepsilon_2$, and in reaction (VII) $\varepsilon_e\ll\varepsilon_{\rm min}$. As a result, gamma-rays with energies of $(5-8.5)\times10^{14}$ eV are generated in the cascade at distances $1.5\leqslant\overline{R}<200$ Mpc from the source. Schemes of cascades generating gamma-rays with energies of $5\times10^{14}-10^{15}$ eV at distances $R\leqslant10$ Mpc from the source were discussed in the earlier publication.[13]

The suggested mechanism of gamma-ray emission accounts for the distribution of gamma-quanta over the celestial sphere reported in Refs. 5–8.

5. INTENSITY OF GAMMA-RAY EMISSION

Now we analyze the intensity of gamma-ray emission in terms of development of electromagnetic cascades in the extragalactic space. Let us assume that extragalactic magnetic fields are weak and electrons generated in cascades do not lose energy in synchrotron radiation.

Let us compare the gamma-ray spectra I_γ detected by the Tien Shan and Bolivian arrays. They are

$$I_\gamma(\geqslant5\times10^{14}\ \text{eV})=(3.4\pm1.1)\times10^{-13}\ (\text{cm}^2\cdot\text{s}\cdot\text{srad})^{-1}[\text{Ref. 6}]$$

and

$$I_\gamma(>10^{14}\ \text{eV})=6.10^{-12}\ (\text{cm}^2\cdot\text{s}\cdot\text{srad})^{-1}[\text{Ref. 8}].$$

The suggested mechanism accounts qualitatively for the difference between measured fluxes[6,8] in the following way. Gamma-rays with an energy of 10^{14} eV have a mean free path of about 30 Mpc in the field of relict photons, and gamma-rays with energies of $(5-8.5)\times10^{14}$ eV have free paths of 10–20 kpc.[23] Therefore, the Bolivian installation detected gamma-rays coming from distances of about 30 Mpc and the Tien Shan array detected gamma-rays created only on the boundaries of the Galaxy, whose number is considerably smaller. (A cascade does not develop inside the Galaxy since electrons lose their energy in synchrotron radiation in galactic magnetic fields.) A more comprehensive analysis of detected fluxes[6,8] requires a detailed comparison between selection criteria for the gamma-ray-initiated showers used by the researchers and computer simulations of electromagnetic cascades in the extragalactic space.

A proton traveling in the extragalactic space interacts with the relict radiation until its energy drops to $E\approx3\times10^{19}$ eV. In each interaction it generates high-energy photons—two photons if reactions (I) and (III) have occurred, and one in reaction (II) with the subsequent chain (IV)–(VI). In our estimates, we assume that $\overline{l}=1.5$ photons are generated in one event. Each photon initiates a cascade developing through reactions (VI) and (VII). Owing to the highly unequal distribution of energy between the particles of a pair in reaction (VII), the cascade has a characteristic feature,[17] namely, one of the particles (electron or photon) has a high energy almost equal to the energy of the initial particle, and the other generated particles have energy $\varepsilon_{e'}\sim\varepsilon_{\rm min}$. Since these particles scatter gamma-ray photons with energies of $5\times10^{14}-10^{15}$ eV over a length of about 5 kpc, let us estimate the number M of such particles on the Galaxy boundary for our analysis of the gamma-ray intensity. We suppose that near the Galaxy there is only one particle in each branch of the cascade and the average number of branches is $\overline{k}$. Then $M=\overline{k}$ and the number of branches $\overline{k}=\overline{m}\,\overline{l}$, where $\overline{m}$ is the average number of proton interactions while the proton energy satisfies $E>3\times10^{19}$ eV. Assuming that the fraction of energy lost by a proton in each interaction is $\overline{K_p}$, we have

$$E(1-\overline{K_p})^m=3\times10^{19}\ \text{eV}.$$

Hence $\overline{m}=5.4$ and $M\approx8$ for $E=10^{20}$ eV and $\overline{K_p}=0.2$. We have set the energy equal to $E=10^{20}$ eV in this estimate because the cross section $\sigma_{p\gamma}$ has a resonance in this region and drops rapidly with decreasing energy.[24]

Possible sources of detected protons with $E>3\times10^{19}$ eV are at distances $R\leqslant R_1=37$ Mpc.[4] Protons generating gamma-ray emission can travel through much longer distances $R_2>R_1$,[13] i.e., their sources are at distances $R\leqslant R_2$ from us. Assuming that the number of sources in a sphere with radius R is $N\sim R^3$, we obtain a relation between the total gamma-ray intensity I_γ and the proton spectrum I:

$$I_\gamma(>10^{14}\ \text{eV})\approx I(>10^{20}\ \text{eV})M(R_2/R_1)^3.$$

The estimate of the distance R_2 follows.

The spectra of cosmic rays in the range $E>10^{20}$ eV measured at different installations vary.[10,27] Let us use in our estimate the integrated intensity[28]

$$I(>10^{20}\ \text{eV})=\begin{pmatrix} & +2\\ 3 & \\ & 1\end{pmatrix}10^{-20}\ (\text{cm}^2\cdot\text{s}\cdot\text{srad})^{-1}$$

and assume that most of the cosmic particles in this energy range are protons. The higher the energy of the initial proton, the larger M, for example, $M=24$ for $E=10^{21}$ eV. Since the particle spectrum is described by a power function and its exponent is about two in the spectral range under discussion,[10,27] the number of particles drops with the energy faster than M grows. Therefore in estimating R_2 we neglect the contribution of protons with $E\gg10^{20}$ eV and take $M=8$. Moreover, we should take into account that in measurements on the Tien Shan array the fraction of the flux from the region $b\geqslant30°$ is 0.625 of the total flux I_γ. Taking the published data on I_γ (Refs. 6 and 8) and I (Ref. 28) within their measurement errors, we obtain $R_2>10^3$ Mpc. Hence sources of protons generating electromagnetic cascades can be at distances of up to several thousands of megaparsecs from us. This estimate of the distance through which protons with energy $E>3\times10^{19}$ eV can travel generating gamma-rays in the interesting energy range is in agreement with calculations[3] and estimates of the previous section.

Let us consider development of an electromagnetic cascade in the energy range $\varepsilon_e<10^{14}$ eV. Such electrons scatter gamma-rays with average energy[10] $\overline{\varepsilon_\gamma}\approx3.6\times10^{-16}\varepsilon_e^2$, and their mean free paths is $\overline{\lambda_e}\sim1$ kpc. The electron energy loss in each scattering is small: $\overline{\Delta\varepsilon_e}/\varepsilon_e\approx10^{-4}$ and 10^{-6} at $\varepsilon_e=10^{12}$ and 10^{10} eV, respectively. In the electron rest frame, the scattering of gamma-ray photons with energy $\varepsilon_\gamma<10^{14}$ eV is almost isotropic.[29] Therefore the transverse distance between two gamma-ray photons sequentially scattered over a length of 1 kpc is ~0.3 kpc, i.e., their density is very small. Therefore detection of cascade photons in the energy range accessible for satellite detectors seems rather difficult.

6. EXTRAGALACTIC MAGNETIC FIELDS

We have assumed that extragalactic magnetic fields are weak. For cascade developing in magnetic fields, electron synchrotron losses should be less than Compton losses. Expressions for these losses are given in Ref. 30. This suggests a limit on the magnetic field strength B:

$$B<(3\times10^{-11}w_b)^{1/2},\quad \varepsilon_e<\varepsilon_{\min},$$

$$B<3.87\times10^{-6}((mc^2)^2/\varepsilon_e\varepsilon_b)$$
$$\times[w_b\ \ln(2\varepsilon_e\varepsilon_b/(mc^2)^2)]^{1/2},\quad \varepsilon_e>\varepsilon_{\min}.$$

Here B is measured in G and the energy density w_b of the background protons is measured in eV/cm^3. From estimates given in Sec. 4 it follows that gamma-ray photons at interesting energies of 5×10^{14}–10^{17} eV are scattered from relict photons by electrons (positrons) with $\varepsilon_e\approx5\times10^{14}$–$10^{17}$ eV; they are scattered from r_1-photons by electrons with $\varepsilon_e\leqslant2\times10^{17}$ eV, and from r_2-photons by electrons with $\varepsilon_e\leqslant10^{16}$ eV. Hence we have $B<2\times10^{-8}$ G if gamma-rays are scattered from relict photons, and $B<2\times10^{-9}$ G if they are scattered from radio photons. Previously some theoretical limits on the magnetic field were published[31]: $B<10^{-9}$ G on the basis of measurements of the quasar rotation measure with $z=2.5$; $B\ll10^{-9}$ G if protons of ultrahigh energies propagate in straight lines in extragalactic space; the ordered magnetic field, if one exists, should satisfy $B<10^{-11}$ G. From these limits it follows that the extragalactic magnetic field does not interfere with electromagnetic cascade development. Since we assumed that protons propagated in extragalactic space in straight lines, $B\ll10^{-9}$ G.

If magnetic fields or their inhomogeneities are higher in some directions than the estimates given above, no gamma-ray-initiated showers should be detected from these directions. This assumption can account for the nonuniformity of the gamma-ray distribution over the celestial sphere. However, unlike the large-scale inhomogeneity in the distribution of active galaxies, the magnitude and type of the extragalactic magnetic field inhomogeneity can hardly be derived from available experimental data.

7. CONCLUSIONS

High-energy gamma-ray emission detected by the Tien Shan and Bolivian arrays[5–8] may result from interaction between extragalactic protons and relict radiation with subsequent generation of electromagnetic cascades in extragalactic space. This mechanism accounts for specific features of this emission, namely, (1) the arrival directions of gamma-rays from the region of predominantly high galactic latitudes $|b|\geqslant30°$; (2) the nonuniformity of the gamma-ray distribution over the celestial sphere detected by the Tien Shan array.

The difference between fluxes measured by the two arrays can be easily explained on a qualitative level. A quantitative analysis requires comparison between selection criteria for showers used by researchers working on the arrays and computer simulations of electromagnetic cascades in extragalactic space.

Cascade development requires that electrons not lose energy to synchrotron radiation in the extragalactic magnetic field. Considering gamma-ray generation, we assumed that protons propagated in extragalactic space along straight lines. From this follows a limit on extragalactic magnetic field strength: $B\ll10^{-9}$ G. This limit does not contradict estimates published previously.[31] The nonuniform distribution of gamma-rays over the celestial sphere can be attributed to the fact that the magnetic field strength or its inhomogeneities satisfy the condition given above only in the region of high galactic latitudes $|b|\geqslant30°$. At present it is difficult to verify this hypothesis in experiment.

In the energy range accessible to devices on satellites, detection of cascade gamma-rays is hardly possible because of their low density.

These conclusions can be verified in experiments on the ANI high-altitude array[32] and on the Tien Shan high-altitude array.[33] Showers will be selected using the technique described by Nikolsky *et al.*[5] Spectra of ultrahigh-energy protons and their arrival directions will be measured by the new EAS-1000 arrays[34] and those described in Refs. 35 and 36, in addition to those listed in Refs. 10 and 27. Moreover, computer simulations of electromagnetic cascades in extragalactic space are required for verification of discussed results.

I am greatly indebted to G. T. Zatsepin for support of this work, to Yu. N. Vetukhnovskaya, V. A. Dogiel, A. V. Zasov, I. G. Mitrofanov, A. I. Nikishov, I. L. Rozental, and O. K. Sil'chenko for discussions. I am also grateful to B. V. Komberg and the referee for constructive criticism.

[1] G. T. Zatsepin and V. A. Kuz'min, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 414 (1966) JETP Lett. [*sic*].
[2] K. Greisen, Phys. Rev. Lett. **16**, 748 (1966).
[3] F. W. Stecker, Phys. Rev. Lett. **21**, 1016 (1968).
[4] A. V. Uryson, JETP Lett. **64**, 77 (1996).
[5] S. I. Nikolsky, Y. N. Stamenov, and S. Z. Ushev, Zh. Éksp. Teor. Fiz. **87**, 18 (1984) [Sov. Phys. JETP **60**, 10 (1984)].
[6] S. I. Nikolsky, J. N. Stamenov, and S. Z. Ushev, J. Phys. G **13**, 883 (1987).
[7] F. Kakimoto, T. Kaneko, H. Yoshii *et al.*, in *Proceedings of 20th ICRC*, Moscow (1987), Vol. **1, p.** 307.
[8] K. Suga, Y. Toyoda, K. Kamata *et al.*, in *Proceedings of the 20th ICRC*, Moscow (1987), Vol. **1, p.** 310.
[9] A. V. Uryson, Kratkie Soobshcheniya Fiz. FIAN, No. 10, 56 (1988).
[10] V. S. Berezinskii, S. V. Bulanov, V. L. Ginzburg, V. A. Dogiel, and V. S. Ptuskin, *Astrophysics of Cosmic Rays*, North-Holland, New York (1990); Nauka, Moscow (1990).
[11] S. Karakula, J. N. Stamenov, and W. Tkaszyk, in *Proceedings of the 19th ICRC*, La Jolla (1985), Vol. **1, p.** 268.
[12] F. A. Aharonian, É. A. Mamidzhanyan, S. I. Nikolsky, and E. I. Tukish, Astrofizika **23**, 55 (1985).
[13] A. V. Uryson, Kratkie Soobshcheniya Fiz. FIAN, No. 3–4, 30 (1996).
[14] J. Rachen, T. Stanev, and P. Biermann, Astron. Astrophys. **273**, 377 (1993).
[15] S. Hayakawa, Prog. Theor. Phys. **37**, 594 (1966).
[16] O. Prilutsky and I. L. Rozental, Acta Phys. Hung. Suppl. 1 **29**, 51 (1970).
[17] L. M. Ozernoi, O. F. Prilutsky, and I. L. Rozental, *Astrophysics of High Energies*, Atomizdat, Moscow (1973).
[18] B. E. Markaryan, V. A. Lipovetsky, and J. A. Stepanyan, Astrofizika **21**, 35 (1984).
[19] V. A. Lipovetsky, S. N. Neizvestny, and O. M. Neizvestnaya, *Soobshcheniya SAO*, No. 55, Siberian Branch of the USSR Academy of Sciences (1987).
[20] M. P. Veron-Cetty and P. Veron, *ESO Scientific Report*, No. 10 (1991).
[21] H. Spinrad, S. Djorgovski, J. Marr *et al.*, PASP **97**, 932 (1985).
[22] H. Kühr, A. Witzel, I. I. K. Pauliny-Toth *et al.*, Astron. Astrophys., Suppl. Ser. **45**, 367 (1981).
[23] J. V. Jelly, Phys. Rev. Lett. **16**, 749 (1966).
[24] Particle Group Data, Phys. Rev. D **50**, Part II (1994).
[25] V. I. Gol'danskiĭ, Yu. P. Nikitin, and I. L. Rozental, *Kinematic Methods in High-Energy Physics* [in Russsian], Nauka, Moscow (1987).
[26] S. Hayakawa, *Cosmic Ray Physics*, Wiley & Sons, N.Y. (1969).
[27] M. N. Dyakonov, T. A. Egorov, N. N. Efimov *et al.*, *Cosmic Radiation of Extremely High Energies* [in Russian], Nauka, Novosibirsk (1991).
[28] G. Brooke, G. Cunningham, P. J. V. Eames *et al.*, in *Proceeding of the 19th ICRC*, La Jolla (1985), Vol. **2**, p. 150.
[29] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, Pergamon Press (1987).
[30] V. L. Ginzburg, *Theoretical Physics and Astrophysics*, Pergamon, Oxford (1979); Nauka, Moscow (1987).
[31] P. P. Kronberg, Rep. Prog. Phys. **57**, 325 (1994); , 809 (1995).
[32] A. Arzumanian, A. A. Chilingarian, V. S. Eganov *et al.*, in *Proceedings of 24th ICRC*, Rome (1995), Vol. **1**, p. 482.
[33] L. I. Vil'danova, P. A. Dyatlov, N. M. Nesterova *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **12**, 79 (1994).
[34] S. S. Ameev, I. V. Chasnikov, Yu. A. Fomin *et al.*, in *Proceedings of 24th ICRC*, Rome (1995), Vol. **1**, p. 466.
[35] M. Teshima *et al.*, Nucl. Phys. B, Proc. Suppl. **28**, 169 (1992).
[36] J. W. Cronin, Nucl. Phys. B, Proc. Suppl. **28**, 213 (1992).

Translation provided by the Russian Editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Fabry–Perot cavity in the field of a gravitational wave

M. V. Sazhin

P. K. Shternberg State Astronomical Institute, 119899 Moscow, Russia

(Submitted 23 April 1997)
Zh. Éksp. Teor. Fiz. **113**, 398–408 (February 1998)

The structure of the electromagnetic field inside a laser cavity—a gravitational-wave detector—is studied. The properties of the spatial and temporal phases of the standing electromagnetic wave are discussed in detail and the corrections appearing in the electric field of the wave as a result of the action of gravitational radiation on the optical system are determined.

[S1063-7761(98)00202-9]

1. INTRODUCTION

Gravitational-wave astronomy is now one of the most rapidly developing fields of science.[1,2] The most promising devices for detecting pulses from space are laser interferometers, which are under development at several centers. The action of a gravitational wave on an interferometer (for example, a Michelson interferometer) can be interpreted as a relative deformation $\Delta l/l \sim h$ of both legs, the deformation being positive in one leg and negative in the other. The dimensionless parameter h is a measure of the deviation of the metric from the Euclidean metric in the field of a gravitational wave.

A realistic prediction of the amplitude of pulses from space suggests a magnitude $h \sim 10^{-21} - 10^{-23}$. Even with a leg of order 10 km displacements $\sim 10^{-17}$ cm must be detected. This is a difficult problem, requiring the application of the principles of quantum nondestructive measurements. Several schemes for perfoming such measurements have been developed. They all have the same drawback—the requirements on the laser pump or the properties of the detecting components are unrealistic.

A QND scheme based on the nonlinear properties of the permittivity of a plate inserted into a microwave circuit was developed recently.[3] The main advantage of this scheme over previous schemes is that the requirements on laser pump power W were sharply reduced, thus making the scheme much more attractive from the standpoint of possible implementation.

The requirements for the measurement procedure and the measuring system itself were likewise formulated in Ref. 3. Specifically, the procedure imposed definite requirements on the structure of the electromagnetic field in the optical cavity. No rigorous substantiation of the response of the system to the action of a gravitational wave was given in Ref. 3. The structure of the field was studied at a qualitative level, where many important details of the process were omitted.

This has made it necessary to analyze in detail the structure of the electromagnetic field inside an interferometer in the field of a gravitational wave. A standing optical wave satisfying the boundary conditions on the mirrors forms in an ideal (no damping) optical cavity. The intensity at the antinodes of the electromagnetic field oscillates with twice the wave frequency and vanishes at the nodes of the field. In ideal cavities a distinction appears between the temporal and spatial phases of the wave: A reference wave is required in order to detect variations of the temporal phase (in gravitational-wave detectors the optical beam propagating in the neighboring leg plays the role of such a wave), while a detector which is decoupled from the mechanical supports and moves in the field of the gravitational wave is required in order to detect variations of the spatial phase, which on account of the boundary conditions is rigidly tied to the mechanical supports of the system. A shift of the nodes of the standing wave inside the cavity is also admissable instead of the shift described above.

The propagation of an electromagnetic wave in the field of a gravitational wave has been studied many times in the literature. It is sufficient to mention several works[4–8] where the propagation of an electromagnetic wave in the field of a gravitational wave was studied in the geometric-optics approximation and under several other simplifying assumptions. In what follows, the propagation of an electromagnetic wave will also be studied in the geometric-optics approximation, and it will also be assumed that the action of the gravitational wave on the interferometer is adiabatic.

In the present paper the structure of an electromagnetic field in ideal cavities with two different configurations will be analyzed—a conventional (linear) Fabry–Perot cavity and an interrupted Fabry–Perot cavity, one leg of which is perpendicular to the other. In the second case the cavity configuration is reminiscent of the configuration of a Michelson cavity. However, there is an important difference. In the Michelson interferometer the legs are coupled by a half-transmitting mirror, whereas in our cavity the corner reflector is an ideally reflecting mirror. In other words, there are no damping or transmission losses in either cavity. However, in a linear Fabry–Perot cavity the position of a node of a standing wave of the electromagnetic field does not vary with the frequency of the gravitational wave, whereas this effect is maximized in the interrupted cavity.

The objective of the present paper is to examine the structure of the electromagnetic field inside each cavity in detail in the geometric-optics approximation, assuming that the action of a gravitational wave is adiabatic. In Sec. 1 the solution of Maxwell's equations in the geometric-optics approximation is studied. In Sec. 2 a simpler case is analyzed—a linear cavity. In Sec. 3 a broken cavity is studied: The structure of the electromagnetic field is determined and the correction introduced in the field by the action of the gravitational wave is calculated.

2. INTERACTION OF LIGHT BEAMS WITH THE FIELD OF A GRAVITATIONAL WAVE

We shall study a gravitational field in the form of a gravitational plane wave given by an equation in a synchronous frame of reference. The metric in the field of such a wave has the form

$$ds^2=(\eta_{\alpha\beta}+h_{\alpha\beta})dx^\alpha dx^\beta,$$

where $\eta_{\alpha\beta}$ is the Minkowski metric, $h_{\alpha\beta}$ are corrections to the metric tensor, and α and β run through the values 0, 1, 2, 3. We assume everywhere below that the speed of light is $c=1$. For simplicity, we shall also assume that the gravitational wave propagates along the z axis and the cavity legs lie in the xy plane. Though limiting somewhat the generality of the analysis, this assumption nonetheless gives the correct qualitative answer to the problem posed. The case of a general position differs from the case studied because additional components of the electromagnetic field arise that complicate the problem. The simplifications preserve the physical meaning of the results without obscuring it with numerous calculations.

In this case only three components of the corrections to the metric are different from zero: $h_{11}=-h_{22},h_{21}$. We shall designate them as $h_{11}=h_+$ and $h_{12}=h_\times$. The correction to the metric can be written in the general form

$$h_{ab}=h_+t_{ab}+h_\times s_{ab}.$$

Here the matrices t_{ab} and s_{ab} describe the tensor structure of the corrections to the metric. For a gravitational wave incident at an arbitrary angle, when the components of the wave vector of the wave in the chosen coordinate system are

$$\mathbf{k}=\Omega(1,\sin\varphi\sin\theta,-\cos\varphi\sin\theta,\cos\theta)$$

(φ is the azimuthal angle and θ is the polar angle of incidence of the wave in the xyz system), the explicit form of the matrices t_{ab} and s_{ab} in terms of the angles of incidence of the gravitational wave is

$$t_{ab}=\begin{pmatrix} \cos^2\varphi-\cos^2\theta\sin^2\varphi & (1+\cos^2\theta)\sin\varphi\cos\varphi & \sin\theta\cos\theta\sin\varphi \\ (1+\cos^2\theta)\sin\varphi\cos\varphi & \sin^2\varphi-\cos^2\theta\cos^2\varphi & -\sin\theta\cos\theta\cos\varphi \\ \sin\theta\cos\theta\sin\varphi & -\sin\theta\cos\theta\cos\varphi & -\sin^2\theta \end{pmatrix}$$

$$s_{ab}=\begin{pmatrix} -\sin 2\varphi\cos\theta & \cos 2\varphi\cos\theta & \cos\varphi\sin\theta \\ \cos 2\varphi\cos\theta & \sin 2\varphi\cos\theta & \sin\varphi\sin\theta \\ \cos\varphi\sin\theta & \sin\varphi\sin\theta & 0 \end{pmatrix}.$$

Now that the general relations have been written out, it is neceassry to state that only one polarization h_+ is studied, though in doing so there is no loss of generality.

We introduce the Fourier representation of the corrections to the metric, omitting everywhere the index $+$:

$$h_{ab}(x^\mu)=t_{ab}\int d\Omega h(\Omega)\exp(i\Omega(t+z)).$$

Strictly speaking, the solution of the problem of the motion of light beams in the field of a gravitational wave presuppose the solution of the generally covariant Maxwell's equations. We take as the components of Maxwell's tensor $E_a=F_{0a}$ and $B_a=-(1/2\sqrt{\gamma})e^{abc}F_{bc}$.[9] These quantities are convenient because their boundary conditions are formulated just as for conventional electric and magnetic field vectors in the theory of electromagnetism: The tangential component of the electric field vanishes at an ideally conducting surface and the normal component of the magnetic field also vanishes on such a surface.

We shall study the solution of the generally covariant Maxwell's equations in the geometric-optics approximation, i.e., we shall assume that the electric field vector is described by the equation (see, for example, Ref. 10)

$$\mathbf{E}=\mathbf{a}e^{-iS}, \tag{1}$$

where **a** is the amplitude of the electric field and S is the eikonal. In the geometric-optics approximation the field vector is represented as a product of a slowly-varying function a and a rapidly varying function e^{-iS}. It should be underscored that the gradient $\partial S/\partial x$ is also a slowly-varying function. Several small parameters arise in the equations. The first one is the amplitude h of the gravitational wave. The second one is the ratio of the characteristic frequency Ω of the gravitational wave to the frequency ω of the electromagnetic wave. We shall perform all calculations to first order in the amplitude of the gravitational wave and to zeroth order in the ratio of the frequencies. Besides these small parameters, which are related to the nature of the gravitational field and the approximation adopted for solving Maxwell's equations, there also exist several small parameters associated with the geometry of the system. We underscore immediately that the magnitudes of the principal small parameters are $h\sim 10^{-23}$ and $\Omega/\omega\sim 10^{-12}$; all other small parameters are much larger. For example, the ratio of the dimensions of the system to the wavelength of the gravitational wave is $\sim 10^{-3}$.

We make use of the slowness of the variations of a and $\partial S/\partial x$ in order to obtain an equation for S:

$$(\eta^{\alpha\beta}+h^{\alpha\beta})\,\frac{\partial S}{\partial x^{\alpha}}\,\frac{\partial S}{\partial x^{\beta}}=0. \tag{2}$$

We solve this equation by the conventional method of expanding in a series in a small parameter—the amplitude h of the gravitational wave—in the form $S=S_0+S_1$. In the zeroth approximation the solution is trivial:

$$S_0=\omega\pi_\alpha x^\alpha,$$

where π_α is a vector that is tangent to the trajectory of the light beam and possesses a zero norm, $\pi_\alpha\pi^\alpha=0$, and ω is the frequency of the electromagnetic wave. A solution of the form

$$\mathbf{E}=\mathbf{a}\,\exp\{-i\omega(t-x)\}$$

describes an electromagnetic plane wave propagating along the x axis.

The first approximation in h has the form

$$\pi^\alpha\,\frac{\partial S_1}{\partial x^\alpha}=\frac{1}{2}\,\omega h_{\alpha\beta}\pi^\alpha\pi^\beta. \tag{3}$$

We now introduce $\tau=t_{\alpha\beta}\pi^\alpha\pi^\beta$. Then the solution of Eq. (3) can be represented as

$$S_1(x^\alpha)=f(\pi_\alpha x^\alpha)+\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\Omega)e^{i\Omega(t+z)}, \tag{4}$$

where f is an arbitrary function of the argument $\pi_\alpha x^\alpha$.

We note that the equation obtained is identical to the equation describing a geodesic curve in the field of a gravitational wave.[7]

The reflection of an electromagnetic wave from the interferometer mirrors in the present formulation of the problem is expressed in the form of boundary conditions for the electromagnetic field on the surfaces of the mirrors. We shall formulate them in the conventional manner.[12] Hence, the boundary conditions must be rigidly tied to the surface of the mirrors. Therefore it is necessary to know the law of motion of the mirrors in the field of the gravitational wave.

In the chosen coordinate system the mirrors move along the lines $x^i=\text{const}$. An analysis of the motion of mirrors in a synchronous frame of reference in the field of a gravitational wave can be found in Ref. 7. For this reason, we shall write our boundary conditions in the form

$$\mathbf{E}_{\text{tang}}(x^i)=0.$$

Besides the approximations indicated above, we shall use one other simplification. We shall assume that the wavelength of the gravitational wave is much longer than the length of both legs $\lambda_{GW}\gg l,L$. On the one hand this assumption greatly simplifies the formulas. On the other hand it makes it possible to neglect transitions between modes of the optical cavity. The effect of a gravitational wave on a cavity containing a standing electromagnetic wave does not just reduce to adiabatic changes in the frequency and phase. A gravitational wave can induce transitions between cavity modes by the parametric interaction of waves of two types.[11] The transition probability is small. This probability contains a small factor, the amplitude of the gravitational wave. Certainly, the adiabatic variation of the frequency and phase is also proportional to the amplitude of the gravitational wave, and in this sense both effects are of the same order of magnitude. However, there is one fundamental detail. Transitions between modes start when the frequency of the gravitational wave is higher than or equal to the frequency difference between neighboring modes:

$$\Omega>\frac{2\pi}{l},$$

where l is the characteristic length of a leg. This inequality implies a limit on the ratio of the wavelength of a gravitational wave and the size of the optical system. If we wish to avoid intermode transitions, then the inequality $\lambda_{GW}>l$ must be satisfied. Therefore, for the condition which we have chosen on the spectral range of sensitivity of the detector, the discussion can be limited to the adiabatic variation of the phase and frequency.

3. LINEAR FABRY–PEROT CAVITY

We now study the structure of the electromagnetic field inside a linear Fabry–Perot cavity. Let the cavity be aligned with the x axis. Let the cavity mirrors be perpendicular to the x axis. Once again, let the gravitational wave propagate along the z axis. We choose the three-dimensional electric field vector in the form $\mathbf{E}=(0,0,E_z)$. This choice serves to simplify the problem and achieves two purposes. First, x and y components of the vector field do not arise in the cavity and it is therefore easier to follow the physical meaning of the calculations. The second advantage is that the boundary conditions for the component E_z on the first and second mirrors are equivalent to the corresponding boundary conditions for the component of the magnetic field in this cavity.

We call the light beam propagating in the positive x direction the first beam and the beam propagating in the negative direction the second beam. Let the first mirror be located at the point $x=0$, $y=0$, $z=0$ and the second mirror at the point $x=l$, $y=0$, $z=0$.

Correspondingly, the field in the first beam will be determined by the equation

$$E^{(1)}=ae^{-iS_1}, \tag{5}$$

$$S_1=\omega(t-x)+\omega\sigma_1+\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\omega)e^{i\Omega t}$$
$$-\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\Omega)e^{i\Omega(t-x)}, \tag{6}$$

and the field in the second beam will be determined by the equation

$$E^{(2)}=-ae^{-iS_2}, \tag{7}$$

$$S_2=\omega(t+x)+\omega\sigma_2+\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}h(\Omega)e^{i\Omega t}$$

$$-\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}h(\Omega)e^{i\Omega(t+x)}. \tag{8}$$

The negative sign in front of the amplitude of the field of the second beam is chosen so as to match the values of the eikonals on the mirrors without a jump in phase by π; and σ_1 and σ_2 are constants determined by the matching conditions. The values of the eikonal in Eq. (1) are real numbers, so that after the integrals in Eqs. (6) and (8) are calculated, their real parts must be taken and substituted into the expressions (5) and (7) for the electric field.

The boundary conditions at the mirrors for the chosen components of the electric field have the form

$$E^{(1)}+E^{(2)}=0.$$

The boundary condition for the first mirror, which is located at the point $x=0$, reduces to the equation

$$\omega(\sigma_1-\sigma_2)=2m\pi,$$

where m is an integer.

The values of the eikonal on the second mirror located at $x=l$ are

$$S_1=\omega(t-l)+\omega\sigma_1+\frac{\omega\tau l}{2}h(t), \tag{9}$$

$$S_2=\omega(t+l)+\omega\sigma_2-\frac{\omega\tau l}{2}h(t), \tag{10}$$

and the boundary condition is expressed in the form of an equation for the eigenfrequencies of the cavity:

$$\omega\left(1-\frac{\tau h}{2}\right)=\frac{m\pi}{l}. \tag{11}$$

In deriving this equation we assumed that the characteristic frequencies of the gravitational wave satisfy the inequality

$$\Omega l\ll 1.$$

Correspondingly, the total electric field inside the cavity is

$$E=a\exp[-i\omega(t+\sigma_2)]\left\{\exp\left[i\omega x\left(1-\frac{\tau h}{2}\right)\right]\right.$$

$$\left.-\exp\left[-i\omega x\left(1-\frac{\tau h}{2}\right)\right]\right\},$$

and the time-averaged squared field is described by an expression that does not depend on the gravitational radiation

$$|E^2|=4a^2\sin^2\frac{m\pi x}{l}.$$

It is evident from the latter formula that the light intensity inside the Fabry–Perot cavity does not depend on the presence or absence of a gravitational wave. In order for the presence of a gravitational wave to influence a quadratic function of the field an additional light source with a standard frequency must be present or the cavity must be nonideal.

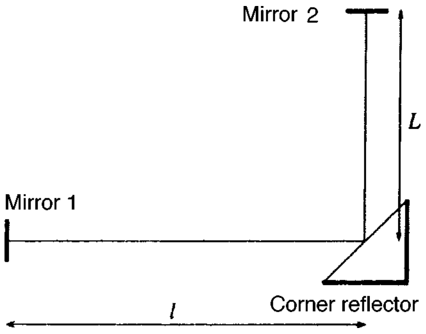


FIG. 1. Arrangement of the optical system. Coordinates of the first mirror $x=0$, $y=-L$. Coordinates of the corner reflector $x=l$, $y=-L$. Coordinates of the second mirror $x=l$, $y=0$. Length of the horizontal leg—l, length of the vertical leg—L.

4. CAVITY WITH PERPENDICULAR LEGS

We now examine the structure of the electromagnetic field inside a Fabry–Perot cavity with perpendicular legs. Let one leg be aligned with the x axis and have length l. Let the cavity mirror be perpendicular to the x axis. Let the second leg be aligned with the y axis and have length L. Let a corner reflector be placed at the intersection of these legs (see Fig. 1). As before, let a gravitational wave propagate along the z axis and the electric field vector be of the form $\mathbf{E}=(0,0,E_z)$. Once again, for this choice of the vector the x and y components of the field do not arise in the cavity. It should also be emphasized that the boundary conditions for the field component E_z at the first and second mirrors are equivalent to the corresponding boundary requirements for the components of the magnetic field in this cavity.

We consider the following path of the beams. At time t_0 a beam leaves the first mirror located at the point $x=0$, $y=-L$, $z=0$, reaches the corner reflector located at the point $x=l$, $y=-L$, $z=0$, reflects from the corner reflector, and propagates toward the second mirror. It reflects from the second mirror, returns to the corner reflector, reflects from the corner reflector, and returns to the first mirror, closing the path. Let the index 1 denote the first part of the path and the index 2 the second part (from the corner reflector to the second mirror). Let the index 3 denote the path of the beam from the second mirror to the corner reflector and, finally, let the index 4 denote the path from the corner reflector to the first mirror. In the calculations we assume Ωl, $\Omega L\ll 1$.

Let the eikonal on each leg and on each path be expressed in the form

$$S_1=\omega(t-x)+\omega\sigma_1+\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}h(\Omega)e^{i\Omega t}$$

$$-\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}h(\Omega)e^{i\Omega(t-x)},$$

$$S_2=\omega(t-y)+\omega\sigma_2-\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\Omega)e^{i\Omega t}$$

$$+\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\Omega)e^{i\Omega(t-y)},$$

$$S_3=\omega(t+y)+\omega\sigma_3-\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\Omega)e^{i\Omega t}$$

$$+\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\Omega)e^{i\Omega(t+y)},$$

$$S_4=\omega(t+x)+\omega\sigma_4+\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\Omega)e^{i\Omega t}$$

$$-\frac{\omega\tau}{2i}\int\frac{d\Omega}{\Omega}\,h(\Omega)e^{i\Omega(t+x)},$$

The boundary conditions on both mirrors for the chosen components of the electric field are, once again, $E^{(1)}+E^{(2)}=0$. For the first mirror, which is located at the point $x=0$, $y=-L$, it reduces to the equation $\omega(\sigma_1-\sigma_4)=2m\pi$, where m is an integer. For the second mirror, located at the point $x=l$, $y=0$, it reduces to the equation $\omega(\sigma_2-\sigma_3)=2n\pi$, where m and n are integers.

The sum of all components of the electric field vanishes on the corner reflector:

$$E^{(1)}+E^{(2)}+E^{(3)}+E^{(4)}=0,$$

and the expressions for each eikonal assume the form

$$S_1=\omega(t+\sigma_1)-\omega l\left(1-\frac{\tau h}{2}\right),$$

$$S_2=\omega(t+\sigma_2)+\omega L\left(1+\frac{\tau h}{2}\right),$$

$$S_3=\omega(t+\sigma_3)-\omega L\left(1+\frac{\tau h}{2}\right),$$

$$S_4=\omega(t+\sigma_4)+\omega l\left(1-\frac{rh}{2}\right).$$

Introducing

$$A=ae^{-i\omega\sigma_1},\quad B=be^{-i\omega\sigma_2}$$

we obtain the boundary condition for the electric components in the form

$$A\sin\omega l\left(1-\frac{\tau h}{2}\right)-B\sin\omega L\left(1+\frac{\tau h}{2}\right)=0. \tag{12}$$

The second boundary condition is imposed for the components of the magnetic field which are perpendicular to the reflecting surface of the corner reflector. We find the magnetic field components themselves from Maxwell's equations, neglecting terms $\sim\Omega/\omega$, i.e., the ratio of the frequency of the gravitational wave to the frequency of the electromagnetic wave. Then, only the components H_y of the magnetic field are different from zero in the first (horizontal) leg and only H_x are different from zero in the second leg. We write out the sums of the magnetic field components in each leg:

$$H_y^{(1)}+H_y^{(4)}=a\left(1-\frac{\tau h}{2}\right)(e^{-iS_1}+e^{-iS_4}).$$

$$H_x^{(2)}+H_x^{(3)}=-b\left(1+\frac{\tau h}{2}\right)(e^{-iS_2}+e^{-iS_3}).$$

The perpendicular component of the magnetic field on the surface of the corner reflector must equal zero, whence we find the equation for the boundary condition

$$H_y^{(1)}+H_y^{(4)}-H_x^{(2)}-H_x^{(3)}=0.$$

This equation reduces to the equation

$$A\left(1-\frac{\tau h}{2}\right)\cos\omega l\left(1-\frac{\tau h}{2}\right)+B\left(1+\frac{\tau h}{2}\right)$$

$$\times\cos\omega L\left(1+\frac{\tau h}{2}\right)=0. \tag{13}$$

Equations (12) and (13) comprise a system of linear equations. Equating the determinant of this system to zero gives a set of characteristic modes of a Fabry–Perot cavity with perpendicular legs. Thus, the solution for the eigenfrequencies of the cavity (to first order in the amplitude of the gravitational wave) is

$$\omega=\frac{m\pi}{l+L}\left(1+\frac{l-L}{l+L}\,\frac{\tau h}{2}\right)+(-1)^{m+1}\frac{\tau h}{2(l+L)}\sin\omega(l-L). \tag{14}$$

We note immediately that the modes of an equal-leg cavity ($l=L$) do not depend on the time or on the presence of the gravitational wave:

$$\omega=\frac{m\pi}{l+L}.$$

When the length of one leg is negligibly small, $L\to 0$, the set of modes is determined by the equation

$$\omega=\frac{m\pi}{l+L}\left(1+\frac{\tau h}{2}\right),$$

i.e., we return to the situation described in the preceding section (one can easily see that the last term in Eq. (14) goes to zero as $L\to 0$).

We shall assume that the legs of our system are symmetric. This greatly simplifies the calculations without loss of generality. Then the eigenfrequencies of the system are constants and do not depend on the presence or absence of a gravitational wave.

Let us calculate the electric field of the first and fourth beams near the corner reflector at some point x located between the first mirror and the corner reflector. The equations for the eikonals in these beams are

$$S_1=\omega(t+\sigma_1)-\omega x\left(1-\frac{\tau h}{2}\right),$$

$$S_4=\omega(t+\sigma_4)+\omega x\left(1-\frac{\tau h}{2}\right).$$

Correspondingly, the sum of the electric fields in the first and fourth beams is

$$E=2a\ \exp[-i\omega(t+\sigma_1+\pi/2)]\sin\ \omega x\left(1-\frac{\tau h}{2}\right),\qquad(15)$$

and the absolute value of the squared field in this point is

$$|E^2|=4a^2\ \sin^2\ \omega x\left(1-\frac{\tau h}{2}\right).$$

It is now obvious that the expression quadratic in the field depends on the gravitational wave. Specifically, the intensity of the radiation inside the cavity with perpendicular legs will depend on the frequency of the gravitational wave.

The constant component of the intensity will equal

$$I=4a^2\ \sin^2\ \omega x,$$

and the variable part, which contains information about the gravitational wave, is

$$\delta I=-4a^2\ \sin(2\omega x)\ \frac{\pi x}{\lambda_e}\ \tau h(t).$$

Here λ_e is the wavelength of the electromagnetic wave. Note that in the case of equal leg lengths ($l=L$) the sum of the first and fourth beams vanishes at the corner reflector. This is a consequence of the symmetry of the legs and the boundary conditions. However, sin $2\omega x$ already reaches a maximum at a distance $\xi=\lambda_e/8$ from the surface of the corner reflector, and the response of the cavity to the gravitational wave is largest there.

Therefore near the corner reflector the quantity quadratic in the electric field oscillates with the frequency of the gravitational wave and with an amplitude of the order of

$$\frac{\pi l}{\lambda_e}\ \tau h(t).$$

5. CONCLUSIONS

In Ref. 3 the necessary condition for implementing the proposed quantum nondestructive meaurement scheme was that the spatial phase of the field (at a point in space rigidly secured to the corner reflector) inside the cavity must contain information about the gravitational radiation. It was shown above that such a situation obtains in a Fabry–Perot cavity with perpendicular legs.

The calculations performed above have a simple physical interpretation. In a linear Fabry–Perot cavity an adiabatic variation of frequency does not produce oscillations of the squared field inside the cavity. In a cavity with perpendicular legs the meaning of the calculations becomes especially transparent in a locally-inertial coordinate system. Indeed, in the limit $\Omega l\ll 1$, in a locally-inertial coordinate system the mirrors move. In the field of a gravitational wave the first mirror moves in a direction toward the corner reflector and the second mirror moves in a direction away from the corner reflector. The sum of the distances between mirror 1 and the corner reflector and between mirror 2 and the corner reflector remains constant during the action of the gravitational wave (for $l=L$), but the position of the corner reflector with respect to each mirror changes. From the standpoint of a one-dimensional observer, who can measure the field only along an axis directed along a light beam, a standing wave with constant characteristics is realized inside the cavity, while the corner reflector and the measuring instrument secured to it ''travel'' along the optical axis. From his point of view the gravitational wave gives rise to oscillations of the corner reflector along the optical axis and therefore the measurement instrument detects motion of the corner reflector according to the change in the field of the standing wave, if, for example, the detector is placed on the ''slope'' of the field of the wave.

I take this opportunity to thank V. B. Braginskiĭ and F. Ya. Khalili for discussions which stimulated the writing of this paper and for numerous valuable remarks.

[1] V. B. Braginskiĭ, Usp. Fiz. Nauk **156**, 93 (1988) [Sov. Phys. Usp. **31**, 836 (1988)].

[2] K. S. Thorne, in *300 Years of Gravitation*, edited by S. W. Hawking and W. Israel, Cambridge University Press, New York, 1987, p. 330.

[3] V. B. Braginsky and F. Ya. Khalily, Phys. Lett. A **218**, 167 (1996).

[4] M. E. Gertsenshteĭn and V. I. Pustovoĭt, Zh. Éksp. Teor. Fiz. **43**, 605 (1962) [Sov. Phys. JETP **16**, 433 (1962)].

[5] R. L. Forward, Phys. Rev. D **17**, 379 (1978).

[6] F. B. Estabrook and H. D. Wahlquest, Gen. Relativ. Gravit. **6**, 439 (1975).

[7] V. N. Rudenko and M. V. Sazhin, Kvant. Élektron. (Moscow) **7**, 2344 (1980) [Sov. J. Quantum Electron. **10**, 1366 (1980)].

[8] V. B. Braginsky, N. S. Kardashev, I. D. Novikov, and A. G. Polnarev, Preprint ICTP IC/89/392 (1989).

[9] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*(Nauka, Moscow, 1988; Pergamon Press, Oxford, 1975).

[10] S. M. Rytov, Yu. A. Kravtsov, and E. I. Tatarskiĭ, *Principles of Statistical Radiophysics, Part II, Random Fields*, (Nauka, Moscow, 1978; Springer, New York, 1987).

[11] V. B. Braginskiĭ, L. P. Grishchuk, A. G. Doroshkevich, Ya. B. Zel'dovich, I. D. Novikov, and M. V. Sazhin, Zh. Éksp. Teor. Fiz. **65**, 1729 (1973) [Sov. Phys. JETP **38**, 865 (1974)].

[12] M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, Oxford, 1969; Nauka, Moscow, 1973).

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Bound states of the muon–antimuon system: lifetimes and hyperfine splitting

S. G. Karshenboĭm*)

D. I. Mendeleev All-Russian Metrological Scientific Research Institute, 198005 St. Petersburg, Russia

V. G. Ivanov

Main Astronomical Observatory, Russian Academy of Sciences, 196140 Pulkovo, Leningrad Region, Russia

U. D. Jentschura and G. Soff

Institut für Theoretische Physik, TU Dresden, 01062 Dresden, Germany

(Submitted 19 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 409–431 (February 1998)

We examine the properties of an atomic system consisting of a muon and antimuon. Expressions are derived for the probability of decay and the hyperfine splitting of the lower levels with allowance for the leading radiative corrections, which are of relative order α. The results for the lifetimes and the ground-state energy are $\tau(1\,{}^3S_1)=1.7907(8)\times10^{-12}$ s, $\tau(1\,{}^1S_0)=0.59547(33)\times10^{-12}$ s, and $E_{\rm hfs}(1s)=4.23284(35)\times10^{7}$ MHz. The relative probabilities for the various decay channels are calculated; in particular, for the $1\,{}^3S_1$ level it is found that $\Gamma(\mu\mu\rightarrow ee\gamma)/\Gamma(\mu\mu\rightarrow ee)\approx15\%$. Finally, possible applications are discussed.
[S1063-7761(98)00302-3]

1. INTRODUCTION

In this paper we study the spectrum of a purely leptonic system consisting of a bound muon and antimuon.[1)] We discuss the properties of the system, such as the lifetimes, decay channels, and the hyperfine splitting of the lower levels, in detail. The given atomic system can be investigated by the standard methods of elementary particle physics. Indeed, dimuonium can be produced in decays and collisions of elementary particles or in nuclear collisions.

Note that atomic decay channels and scattering processes have already been observed. For instance, the bound $\pi\mu$ system was detected in the $K_L^0\rightarrow(\pi\mu\text{-atom})+\nu$ decay by Coombes *et al.*,[1] and then Aronson *et al.*[2] were able to increase the number of events and carried out more accurate measurements. Ultrarelativistic positronium in the final state was detected in proton–carbon collisions.[3] The decay $\pi^0\rightarrow\mathrm{Ps}+\gamma$ inside the target was assumed to be the main source of the atoms, and additional experiments[4] proved that this is indeed the case. Recently experiments with a proton beam and a tantalum target produced pionium.[5]

To make our description complete, we mention the detection of the Coulomb interaction in the final state of proton–proton scattering.[6] In the latter case the interaction was observed in the $\pi^+\pi^-$, $p\pi^-$, and K^+K^- channels, although given the small number of events, it was impossible to determine whether these states belong to the continuous or discrete spectrum.

The production of bound states of the $\mu^+\mu^-$ system is even a rarer event than the production of hadronic atoms, so that it is difficult to expect that enough atoms will be produced for spectroscopic studies. Usually the production probability is proportional to the wave function at the origin, i.e., to δ_{l0}/n^3. Clearly, in this case we can only hope to obtain the lowest s states. Their lifetimes are in the picosecond (10^{-12} s) range, i.e., they are of the same order as the lifetimes of some neutral mesons (cf., e.g.,[7] $\tau(K_L^0)=89.3(1)\times10^{-12}$ s, $\tau(D^0)=0.415(4)\times10^{-12}$ s, $\tau(B^0)=1.56(6)\times10^{-12}$ s, and $\tau(B_S^0)=1.61(10)\times10^{-12}$ s), which have been measured to within about one percent. The decay products are also characteristic of the decay of the particles: the principal mode for paradimuonium is annihilation into two gamma photons (Fig. 1a), while for orthodimuonium conversion into an electron–positron pair via virtual one-photon annihilation (Fig. 1b) is the principal mode. In both cases the energy of the final particle in the atom's center-of-mass system is 106 MeV.[2)]

Recent progress in observing various atomic states in collisions[3,5] and decays,[1,2,4] and the characteristic decay channels of dimuonium, suggest that dimuonium will be detected soon. From the standpoint of experimental high-energy physics, dimuonium is a family of neutral scalar (paradimuonium, states $1s$ and $2s$) and vector (orthodimuonium, states $1s$ and $2s$) bosons. The study of dimuonic atomic states is of interest to experimentalists as a model for the search for exotic neutral particles with masses and lifetimes characteristic of ordinary particles, but with anomalously weak interaction between the two types. Note that many models have a place for new neutral particles, either extremely heavy and interacting with matter, or with a normal mass but weakly interacting with ordinary particles. Clearly, dimuonium has the right properties for neutral exotic particles of the second kind. Allowing for the fact that lately the search for such particles has become one of the main lines of experimental research, we believe that the aforementioned properties of dimuonium make it a good test particle.

The $\mu^+\mu^+$ system may also be of interest as a test of

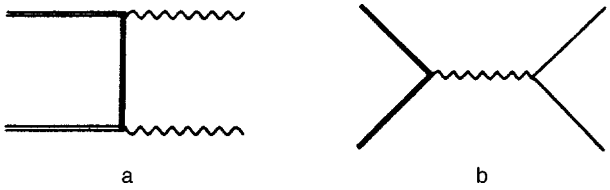


FIG. 1. Decay of dimuonium in the leading approximation: (a) paradimuonium, and (b) orthodimuonium. Here and in what follows we depict only one diagram of each type. The double fermion line corresponds to the muon and the single line, to the electron. For instance, in the case of diagram (a) one must allow (in a symmetric way) for two diagrams with different photons.

quantum electrodynamics (QED). First, as in the case of ordinary muonic atoms, it is possible to check the predictions of QED at spacelike momenta of the order of the electron mass (a typical atomic momentum is $\alpha m_\mu/2 \simeq 0.75 m_e$). In ordinary muonic atoms, effects related to nuclear structure play an important role (these effects are not present in a purely leptonic system). The expected relatively modest accuracy of measured lifetimes and hyperfine splitting may be partly compensated by corrections to the wave functions that are more sensitive to the features of the potentials than are such integrated quantities as the energy. On the other hand, in calculating the hyperfine splitting (Fig. 2b) and the decay width (Fig. 1b) for the orthodimuonic system, we must allow for the annihilation diagram, whose contribution is proportional to the photon propagator. Measuring the lifetime and hyperfine splitting in orthodimuonium, we can experimentally study vacuum polarization at a high timelike momentum (in units of electron mass) that is still below the pionic threshold.

In our opinion, the most interesting quantities are lifetimes; finding these to within about one percent requires knowing the radiative corrections and the ratios of the partial widths of the various decay channels, which can be directly used in the detection process. Total and partial widths are discussed in Secs. 3 and 4. But before that, in Sec. 2, we study hyperfine splitting of the $1s$ and $2s$ states. The point is that on the one hand, parts of the diagrams for hyperfine splitting and decay of orthodimuonium resemble one another, while on the other, there is an analogy between hyperfine splitting in dimuonium and positronium, which has been well studied. The results are discussed at the end of the paper.

Before proceeding, we briefly review some familiar results obtained for exotic atoms. Questions associated with the production, spectrum, and decay of pionium have been studied most thoroughly (see Refs. 8 and 9 and the literature cited therein). Other exotic atoms containing mesons have also been studied (see, e.g., Refs. 8, 10, and 11). Only a few papers have been devoted to the study of dimuonium proper. In Ref. 10, in addition to other atomic states participating in the decay of elementary particles, the process $\eta \to (\mu^+\mu^-\text{-atom}) + \gamma$ was investigated. Bilen'kiĭ *et al.*[8] studied processes with dimuonium production in the reaction $\pi^- + p \to (\mu^+\mu^-\text{-atom}) + n$ and in the scattering of a photon in the field of a nucleus. Several radiative corrections to the decay of heavy leptoniums in the ortho-state were investigated by Malefant,[12] who found only some of the contributions of relative order α (see Sec. 3 for more details). At the same time it must be noted that some of the results for hyperfine splitting in positronium[13,14] can be used in the case of dimuonium and to calculate the orthodimuonium decay width.

2. HYPERFINE SPLITTING IN DIMUONIUM

It is convenient to start the discussion of hyperfine splitting in dimuonium with the results obtained for positronium. The leading contribution is of order $\alpha^4 m$ and stems from two diagrams in Fig. 2: the exchange of a transverse photon (a) and virtual annihilation (b). The result for dimuonium differs in that the electron mass must be replaced by the muon mass:

$$E^{(0)}_{\rm hfs}(ns) = \frac{E_F}{n^3}, \tag{1}$$

where the Fermi energy E_F is defined as follows[3)]:

$$E_F \simeq \tfrac{7}{12}\alpha^4 m_\mu \simeq 0.175 \text{ eV} = 4.23\times 10^7 \text{ MHz}. \tag{2}$$

The factor 7/12 results from adding 1/3 (transverse exchange) to 1/4 (annihilation).

The leading radiative corrections of order $\alpha^5 m$ stem from the diagrams depicted in Fig. 3, where the contributions from diagrams a to e have the same form for positronium and dimuonium, while the contributions from diagrams f to j are specific to dimuonium. The results for the hyperfine splitting of the ground state in positronium were found by Karplus and Klein,[13] and they were later corroborated by Fulton and Martin,[14] who established that for the ns levels the entire dependence on the state reduces to a scaling factor n^{-3}, which emerges from the square of the Schrödinger wave function at the origin, $|\varphi_{ns}(0)|^2$. The result is[13,14]

$$\Delta E^{\rm Ps}(ns) = \frac{\alpha}{\pi}\left[-\frac{32}{21} - \frac{6}{7}\ln 2 + \frac{3}{7}\pi\ i\right]\frac{E_F}{n^3}. \tag{3}$$

As noted earlier, this result stems from the first five diagrams in Fig. 3, which include the anomalous magnetic moment (a), two-photon exchange (b), vertex correction to annihilation (c), polarization insertion into an annihilation photon (d), and two-photon annihilation (e). All these contributions are listed separately in Table I. For the discussion that follows, we present two of these corrections in detail: the vertex correction

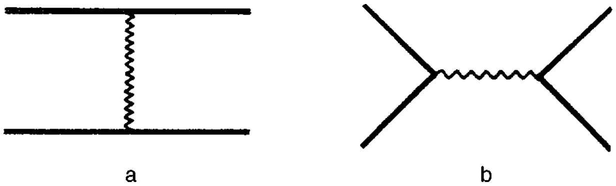


FIG. 2. Hyperfine splitting in dimuonium in the leading approximation: the ''spring'' denotes the exchange of a transverse photon (a), and diagram (b) depicts virtual one-photon annihilation (the wavy line denotes an ordinary photon).

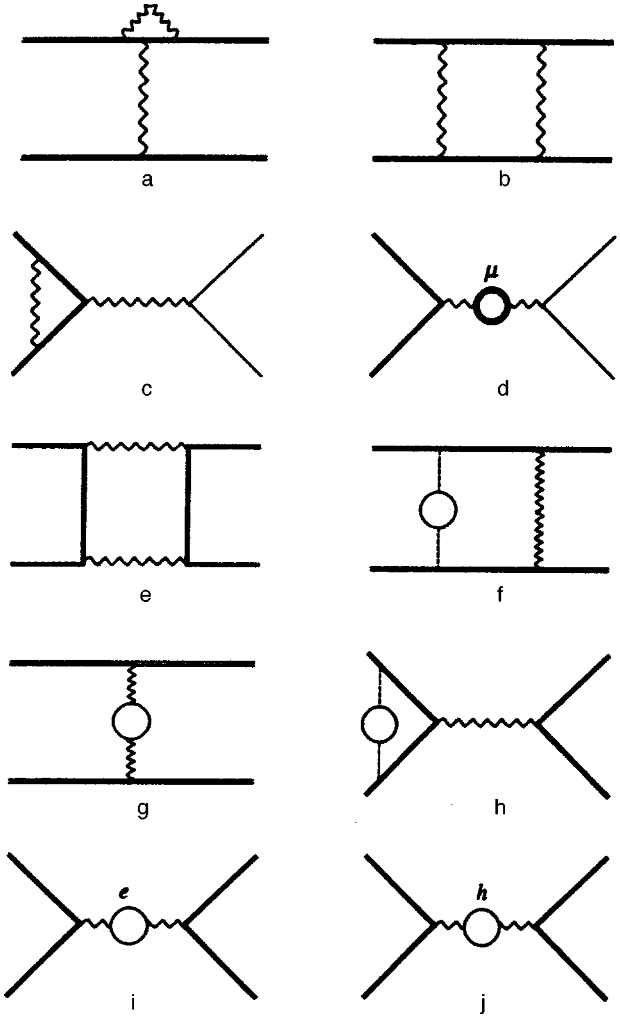


FIG. 3. Single-loop radiative corrections to hyperfine splitting in dimuonium: (a) contribution of the anomalous magnetic moment, (b) two-photon exchange, (c) the vertex correction to annihilation, (d) muonic polarization of vacuum, (e) virtual two-photon annihilation, (f) transverse exchange with the Uehling correction to the wave function (the Coulomb electron is depicted by a dashed line), (g) vacuum polarization in a transverse photon, (h) one-photon annihilation with a correction to the wave function, (i) electron vacuum polarization in an annihilation photon, and (j) hadronic vacuum polarization. In some cases (b and c) one must subtract the contributions of the preceding order.

$$\Delta E^{\mathrm{V}}(ns)=\frac{\alpha}{\pi}[-2\times 2]\,\frac{3}{7}\,\frac{E_F}{n^3} \tag{4}$$

(here we have written the combinatorial factor 2 explicitly), and the insertion of vacuum polarization from the leptons comprising the atom (i.e., electrons for positronium and muons for dimuonium),

$$\Delta E^{\mathrm{P}\mu}(ns)=\frac{\alpha}{\pi}\left[-\frac{8}{9}\right]\frac{3}{7}\,\frac{E_F}{n^3}. \tag{5}$$

The sign of the imaginary part in (3) can be explained by the fact that hyperfine splitting is the difference between the orthopositronium and parapositronium energies, with the imaginary part giving the probability of parapositronium decay (due to two-photon annihilation)[15]:

$$\Gamma^{(0)}_{\mathrm{PM}}(ns)=\frac{\alpha^5 m_\mu}{2n^3}. \tag{6}$$

The corresponding lifetime is $\tau^{(0)}_{\mathrm{PM}}(ns)=n^3\times 0.6021\times 10^{-12}$ s.

We begin the discussion of the specific contributions with effects related to the Uehling potential. As noted earlier, the leading contribution (1) is proportional to $|\varphi_{ns}(0)|^2$, which varies due to corrections to the potential. Vacuum polarization in a muonic atom, with a small nuclear charge, is a nonrelativistic effect.

The insertion of vacuum polarization induces the following change in the photon propagator:

$$\frac{1}{q^2+i0}\rightarrow\frac{1}{q^2+i0}I_P(q^2)=\frac{\alpha}{\pi}\int_{s_0}^{\infty}ds\rho(s)\,\frac{1}{q^2-s+i0}. \tag{7}$$

Strictly speaking, the insertion of vacuum polarization is always transverse, i.e., proportional to $g_{\mu\nu}-q_\mu q_\nu/q^2$, with the result that (7) is true only in the Landau gauge. It is clear, however, that by appropriately selecting the longitudinal part of the photon propagator and adding terms of order α we can ensure the validity of (7) in any covariant gauge (for details see Ref. 16).

In the case of electron polarization the parameter $s=\lambda^2$ and the integral of the spectral function $\rho(s)$ can be represented in the form

$$\lambda=\frac{2m_e}{\sqrt{1-v^2}} \tag{8}$$

and

$$I_{\mathrm{Pe}}(q^2)=\frac{\alpha}{\pi}\int_0^1 dv\,\frac{v^2(1-v^2/3)}{1-v^2}\,\frac{q^2}{q^2-\lambda^2+i0}. \tag{9}$$

As a result, the polarization insertion into the Coulomb photon is described by the nonrelativistic potential

$$V_U(\mathbf{r})=-\frac{\alpha}{\pi}\int_0^1 dv\,\frac{v^2(1-v^2/3)}{1-v^2}\,\frac{\alpha e^{-\lambda r}}{r}. \tag{10}$$

The Uehling potential (10) leads to a shift in the energy levels (the Lamb shift) and to corrections to the wave function. The shift in the energy levels,

$$\mathscr{E}(1s)=-\frac{\alpha}{\pi}[0.15\ldots]E_0=-0.49\ \mathrm{eV},$$

$$\mathscr{E}(2s)=-\frac{\alpha}{\pi}[0.072\ldots]\,\frac{E_0}{4}=-0.058\ \mathrm{eV},$$

$$\mathscr{E}(2P)=-\frac{\alpha}{\pi}[0.00172\ldots]\,\frac{E_0}{4}=-0.0014\ \mathrm{eV} \tag{11}$$

must be compared with the analog of the Rydberg constant for dimuonium

$$E_0=\frac{\alpha^2 m_\mu}{4}=1406.6133(5)\ \mathrm{eV},$$

which was calculated using the following values of the physical constants: $\alpha^{-1}=137.035\,9895(61)$ and $m_\mu=105.658389(91)$ eV (see Ref. 7).

Clearly, the contributions corresponding to diagrams f and h in Fig. 3 can be found by taking the perturbation of $|\varphi_{ns}(0)|^2$ into account, which enters directly into the leading contributions (Fig. 2). The correction is of relative order α and depends on the state (cf. Ref. 17). The desired quantity has the form

$$\frac{\Delta|\varphi_{ns}(0)|^2}{|\varphi_{ns}(0)|^2}=2\int d^3r\overline{G}_{ns}(0,\mathbf{r};E_{ns})V_U(\mathbf{r})\varphi_{ns}(\mathbf{r}), \qquad (12)$$

where

$$\overline{G}_{ns}(E_{ns})=\sum_{n'\neq n}\frac{|\psi_{n's}\rangle\langle\psi_{n's}|}{E_{ns}-E_{n's}} \qquad (13)$$

is the reduced nonrelativistic Coulomb Green's function (more precisely, its s-wave part). The simplest way to calculate the integral in (13) is to use the explicit expression for the Green's function (see, e.g., Ref. 18):

$$\overline{G}_{1s}(E_{1s};0,\mathbf{r})=\frac{\alpha m_r^2}{4\pi}\frac{e^{-\rho}}{\rho}[4\rho(\ln 2\rho+C)+4\rho^2-10\rho-2] \qquad (14)$$

and

$$\overline{G}_{2s}(E_{2s};0,\mathbf{r})=-\frac{\alpha m_r^2}{4\pi}\frac{e^{-\rho/2}}{2\rho}[4\rho(\rho-2)(\ln\rho+C)+\rho^3-13\rho^2+6\rho+4], \qquad (15)$$

where we have introduced the dimensionless radius $\rho=\alpha m_r r$, and $C=0.5772\ldots$ is Euler's constant. We also allowed for the fact that the nonrelativistic problem can always be solved in terms of the reduced mass m_r.

Now it is easy to calculate the integral with respect to the radius analytically and to integrate the result with respect to the auxiliary parameter v numerically. The results are

$$\frac{\Delta|\varphi_{1s}(0)|^2}{|\varphi_{1s}(0)|^2}=\frac{\alpha}{\pi}[1.059\ldots] \qquad (16)$$

and

$$\frac{\Delta|\varphi_{2s}(0)|^2}{|\varphi_{2s}(0)|^2}=\frac{\alpha}{\pi}[0.916\ldots]. \qquad (17)$$

Vacuum polarization usually leads to small numerical contributions, which can easily be seen in the case of Lamb shifts (11). But for wave functions at the origin the contributions prove to be relatively large. Note that wave functions are much more sensitive to the details of the potential. What is more important, however, is that a particular value of the wave function is not an integrated quantity, as is the energy (11) or the matrix element for the atomic transition probability,[17] where the corrections are of a normal order of magnitude. The shape of the potential can have a strong effect on the wave function.

In order to understand the nature of the large numerical values of the contributions, we write the Green's function as a sum over states,

$$\Delta\varphi_{ns}(0)=\int d^3r\overline{G}_{ns}(E_{ns};0,\mathbf{r})V_U(r)\varphi_{ns}(\mathbf{r})=\int d^3r\sum_{n\neq n'}\frac{\varphi_{n's}(0)\varphi^*_{n's}(\mathbf{r})}{E_{ns}-E_{n's}}V_U(\mathbf{r})\varphi_{ns}(\mathbf{r}), \qquad (18)$$

and ignore the energy of the intermediate state n' in the denominator $E_{ns}-E_{n's}$. Clearly, we can easily find the sum in (18) and the result, with $E_{ns}=E_{n's}$ replaced by E_{ns}, proves to be divergent,

$$\frac{V_U(0)-\langle ns|V_U|ns\rangle}{E_{ns}}\varphi_{ns}(0).$$

Note that the Uehling potential $V_U(r)$ (Eq. (10)) is inversely proportional to the radius, and the weighting function is logarithmically divergent.

This divergence means that we cannot ignore the intermediate state energy. On the other hand, it is clear that for the $1s$ level, for example, in estimating the contribution of the discrete spectrum we are always justified in ignoring this energy. This suggests that the continuum states provide a considerable contribution, with the sum over states slowly decreasing with increasing wave number. Indeed, explicit calculations show that the high-energy continuum states are responsible for the large numerical value (the first term in square brackets corresponds to the discrete spectrum and the second, to the continuum):

$$\frac{\Delta|\varphi_{1s}(0)|^2}{|\varphi_{1s}(0)|^2}=\frac{\alpha}{\pi}[0.07+0.98] \qquad (19)$$

and

$$\frac{\Delta|\varphi_{2s}(0)|^2}{|\varphi_{2s}(0)|^2}=\frac{\alpha}{\pi}[-0.26+1.17]. \qquad (20)$$

Allowance for the correction to the wave function immediately leads to the contributions of diagrams f and h in Fig. 3 (see Table I). Note that the results depend on the state. Another state-dependent contribution stems from the vacuum polarization in the transverse photon (diagram g in Fig. 3). The corresponding contribution is proportional to

$$\mathscr{M}_{nl}=\langle nl|\nabla^2V_U|nl\rangle. \qquad (21)$$

The integral with respect to the radius can easily be calculated analytically, and the integral with respect to the parameter v, numerically. The results for the $1s$ and $2s$ levels are listed in Table I.

Note that state-dependent contributions are more sensitive to the potential than the Lamb shift. In the case of corrections to the wave function this manifests itself in the predominance of high-frequency intermediate states, while for vacuum polarization in a transverse photon the contribution is proportional to the derivative of the potential.

Now we calculate the contributions related to the insertion of vacuum polarization into an annihilation photon. We

TABLE I. Corrections of order $\alpha^5 m$ to the hyperfine splitting of the levels $1s$ and $2s$ are given in relative units: $\Delta E(ns)=(\alpha/\pi)CE_F/n^3$. The notation of the contributions corresponds to Fig. 3. The contributions a to e are the same as in positronium, and the contributions f to j are specific to dimuonium.

Contribution	$C(1s)$	$C(2s)$
a	0.571	0.571
b	−0.857	−0.857
c	−1.714	−1.714
d	−0.381	−0.381
e	0.263	0.263
f	0.605	0.523
g	0.345	0.355
h	0.454	0.393
i	1.483	1.483
j	−0.080(9)	−0.080(9)
Total	0.689(9)	0.556(9)

found the contribution of muonic polarization above, so that we need to examine electron polarization (diagram i in Fig. 3) and hadronic polarization (diagram j in Fig. 3). Electron vacuum polarization can easily be taken into account if we use the well-known asymptotic behavior of the polarization operator at high momenta:

$$I_{\mathrm{Pe}}(q^2)=\frac{\alpha}{\pi}\left[\frac{1}{3}\ln\frac{-q^2}{m_e^2}-\frac{5}{9}\right]=\frac{\alpha}{\pi}\left[\frac{1}{3}\ln\frac{q^2}{m_e^2}-\frac{5}{9}-\frac{\pi}{3}i\right]. \tag{22}$$

The corresponding contribution is $(q^2=(2m_\mu)^2)$

$$\Delta E_{\mathrm{Pe}}(ns)=\frac{3}{7}\,\frac{\alpha}{\pi}\left[\frac{1}{3}\ln\frac{(2m_\mu)^2}{m_e^2}-\frac{5}{9}-\frac{\pi}{3}i\right]\frac{E_F}{n^3}. \tag{23}$$

The presence of an imaginary part means that orthodimuonium has a finite lifetime, determined by the decay $(\mu^+\mu^-\text{-atom})\rightarrow e^+e^-$ (Fig. 1b):

$$\Gamma^{(0)}_{\mathrm{OM}}(ns)=\frac{\alpha^5 m_\mu}{6n^3}. \tag{24}$$

The lifetime is $n^3\times1.806\times10^{-12}$ s (see Ref. 8). Here we note an important difference between positronium and dimuonium: the lifetimes of the para- and ortho-states in dimuonium are of the same order (cf. Eqs. (6) and (24); see Fig. 1), while orthopositronium (Fig. 4b), which decays into three gamma photons,[19] lives much longer than parapositronium, which decays into two photons (see Eq. (6) and Fig. 4a). The difference appears because dimuonium has an additional mode of decay into an electron–positron pair, which proves to be the dominant one (Eq. (24)).

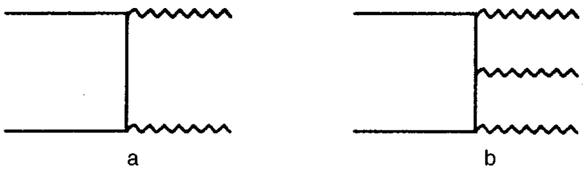


FIG. 4. Decay of positronium in the leading approximation: (a) parapositronium, and (b) orthopositronium.

Now let us discuss the hadronic contribution to vacuum polarization. We first examine the approximation used in Ref. 20 to calculate the contribution of hadronic polarization to hyperfine splitting in muonium. The method consists in the following. The spectral function in the dispersion integral (7) is approximated by the sum of the following terms: the pionic contribution found by using the form factor of Ref. 21, the contributions of the ω- and ϕ-mesons in the pole approximation, and the background above 1 GeV. The parameters needed for the calculations (particle masses, coupling constants, parameters that contribute to the form factor, and parameters characterizing the background) were taken from Refs. 7 and 22. A more detailed discussion can be found in Appendix A; the final results for hadronic polarization are listed in Table I.

Table I also lists the contributions of individual terms to the hyperfine splitting. The final result for the correction is

$$\Delta E_{\mathrm{hfs}}(1s)=\frac{\alpha}{\pi}\,0.689(9)E_F \tag{25}$$

and

$$\Delta E_{\mathrm{hfs}}(2s)=\frac{\alpha}{\pi}\,0.556(9)\,\frac{E_F}{8}. \tag{26}$$

3. LIFETIME OF ORTHODIMUONIUM

We found the leading contribution (see Fig. 1b) earlier:

$$\Gamma^{(0)}_{\mathrm{OM}}(ns)=\frac{\alpha^5 m_\mu}{6n^3}.$$

As noted earlier, there is a close analogy between a number of contributions of single-loop corrections to hyperfine splitting and to the orthodimuonium decay width, whose amplitude is depicted diagrammatically in Fig. 5. Clearly, in the given approximation there are three decay channels for orthodimuonium, with the final products ee, $ee\gamma$, and 3γ. We begin with conversion into an electron–positron pair (diagrams a to f in Fig. 5). The main results are listed in Table II. Note that these results are given for the width, which incorporates the square of the amplitudes, so the correction is doubled. The first five corrections (diagrams a to e in Fig. 5) yield equal contributions (in relative units) both to the annihilation part of the hyperfine splitting (which amounts to 3/7 of the total contribution) and to the orthodimuonium decay width. However, there may be different combinatorial factors. For the width, all combinatorial factors are equal to two, while for the hyperfine splitting, no contributions except the vertex are doubled.

The calculation of diagram f in Fig. 5 leads to an infrared divergence, which cancels diagram g in Fig. 5, corresponding to another decay channel (OM$\rightarrow ee\gamma$). The total decay probability, which determines the lifetime of the state, incorporates the IR finite sum of the diagrams f and g in Fig. 5, so that we first find the total contribution. We seek the width as the imaginary part of the energy. Here the leading contribution corresponds to diagram a in Fig. 6, and the desired correction to diagram b in Fig. 6. Earlier, in discussing the contribution of electron polarization to hyperfine splitting

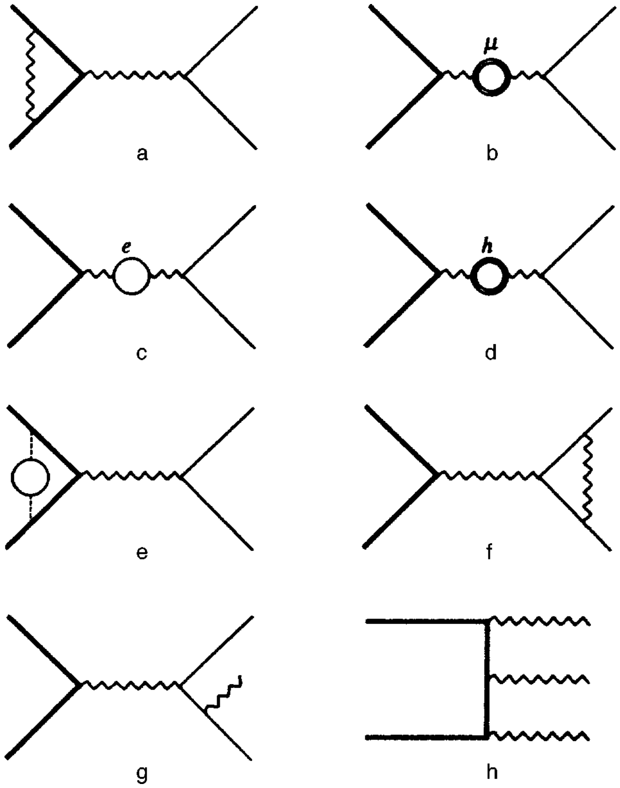


FIG. 5. Single-loop radiative corrections to the orthodimuonium decay amplitude: (a) vertex correction for the annihilation of the bound muonic pair, (b) muonic vacuum polarization, (c) electron vacuum polarization, (d) hadronic vacuum polarization in virtual annihilation, (e) virtual annihilation with the Uehling correction to the wave function, (f) vertex correction for the creation of a free electron pair, (g) pair production with bremsstrahlung, and (h) three-photon annihilation. In calculating diagrams a one must subtract the contribution of the preceding order.

(see Eqs. (22) and (23)), we noted that the imaginary part is completely defined by the coefficient of the logarithm. The asymptotic behavior of two-loop vacuum polarization at high momenta is well known (see, e.g., Ref. 23):

$$I^{(2)}_{\mathrm{Pe}}(q^2)=\left(\frac{\alpha}{\pi}\right)^2\left[\frac{1}{4}\ln\frac{-q^2}{m_e^2}+\left(\zeta(3)-\frac{5}{24}\right)\right]$$

$$=\left(\frac{\alpha}{\pi}\right)^2\left[\frac{1}{4}\ln\frac{q^2}{m_e^2}+\left(\zeta(3)-\frac{5}{24}\right)-\frac{\pi}{4}i\right], \qquad (27)$$

where $\zeta(3)=1.2020569\ldots$ is the Riemann zeta function, and of course in the calculations we must put $q^2=4m_\mu^2$. Clearly, the relative value of the correction is exactly equal to the ratio of the coefficients of the logarithms in (22) and (27), i.e., $(3/4)\alpha/\pi$.

TABLE II. Corrections of order $\alpha^6 m$ to the decay width of the $1s$ and $2s$ levels in orthodimuonium are given in relative units: $\Delta\Gamma_{\mathrm{OM}}(ns)=(\alpha/\pi)C\Gamma^{(0)}_{\mathrm{OM}}(ns)$. The notation of the contributions corresponds to Fig. 5.

a	−4.00	−4.00
b	6.92	6.92
c	−1.78	−1.78
d	−0.37(4)	−0.37(4)
e	1.06	0.92
f+g	0.75	0.75
h	1.16	1.16
Total	3.74(4)	3.60(4)

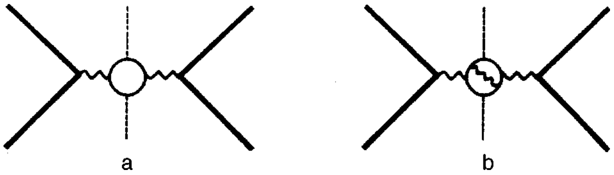


FIG. 6. Some diagrams for the width of orthodimuonium decay as the imaginary part of the energy: (a) leading contribution, and (b) allowance for a free electron vertex and bremsstrahlung.

The partial widths corresponding to individual decay modes incorporate infrared-divergent contributions separately. To ensure their finiteness we must add to the vertex contribution the probability of soft-photon bremsstrahlung (for more details see Ref. 24). Assuming that all the characteristic momenta in the final state are of the order of the muon mass and are known to 1%, we can conclude that the photon cutoff frequency, which separates the detected photons from the soft, is of the order of the electron mass, $\Delta\omega\sim m_e$. In this case it is sufficient to find the leading contribution in the doubly logarithmic approximation, since the first power of the logarithm already depends on the detection specifics for photons with frequency of order m_e, yielding a result that is excessively accurate. An expression for the vertex function can be found in Ref. 24, and the contribution of bremsstrahlung differs only in sign:

$$\Delta\Gamma^{ee\gamma}_{\mathrm{OM}}(ns)=2\,\frac{\alpha}{\pi}\ln^2\frac{m_\mu}{m_e}\Gamma^{(0)}_{\mathrm{OM}}(ns), \qquad (28)$$

which amounts to approximately 13%. The accuracy of this calculation is determined by unknown contributions, with the power of the logarithm being one unit smaller ($\ln(m_\mu/m_e)\approx 5.3$). Usually the numerical coefficients of the nonleading logarithms are larger than those of the leading logarithms, and we estimate the error to be 30% of (28).

The width of three-photon decay in orthopositronium found in Ref. 19, after we have replaced the electron mass by the muon mass, becomes

$$\Delta\Gamma^{3\gamma}_{\mathrm{OM}}(ns)=\frac{\alpha}{\pi}\frac{4}{3}(\pi^2-9)\Gamma^{(0)}_{\mathrm{OM}}(ns). \qquad (29)$$

The results for the correction of order $\alpha^6 m$ to the decay width are listed in Table II. The final expression is

$$\Delta\Gamma_{\mathrm{OM}}(1s)\simeq\frac{\alpha}{\pi}\{1.90+0.68(4)+1.16\}\Gamma^{(0)}_{\mathrm{OM}}(1s)$$

$$=\frac{\alpha}{\pi}\,3.74(4)\Gamma^{(0)}_{\mathrm{OM}}(1s). \qquad (30)$$

Here we have identified three types of contribution. The first term in braces corresponds to the contributions of diagrams a to c and f in Fig. 5,

$$\left(\frac{4}{3}\ln\frac{2m_\mu}{m_e}-\frac{221}{36}\right),$$

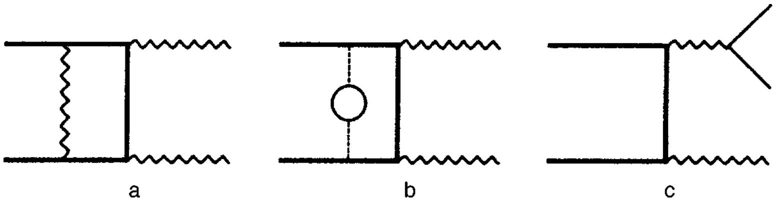


FIG. 7. Radiative corrections of relative order α to the paradimuonium decay amplitude: (a) single-loop correction to the annihilation of a bound muonic pair, (b) two-photon annihilation with the Uehling correction to the wave function, and (c) electron–positron pair production. In calculating diagrams a one must subtract the contribution of the preceding order.

obtained analytically; the second term corresponds to the contributions of diagrams d and e in Fig. 5 obtained numerically; and the third term corresponds to three-photon exchange. As noted in the Introduction, there was an earlier attempt to determine the radiative corrections,[12] but only the first (analytic) contribution was found, and this amounted to only half of the final result. As for the second contribution, only the Uehling correction to the wave function was taken into account, and in calculating this correction, Malefant[12] used the free Green's functions of the muon and antimuon in the initial expression without allowing for the Coulomb interaction, which of course led to incorrect results. The contribution of hadronic vacuum polarization was ignored completely (as was the third term in braces). We also note that the only state-dependent contribution is the second (numerical) one. For instance, in the case of the $2s$ level we have

$$\Delta\Gamma_{\rm OM}(2s)\simeq\frac{\alpha}{\pi}\{1.90+0.54(4)+1.16\}\Gamma^{(0)}_{\rm OM}(2s)$$

$$=\frac{\alpha}{\pi}\,3.60(4)\Gamma^{(0)}_{\rm OM}(2s). \tag{31}$$

4. LIFETIME OF PARADIMUONIUM

We now discuss the decay of paradimuonium. The leading contribution to its width is similar to the corresponding contribution for parapositronium[15]:

$$\Gamma^{(0)}_{\rm PM}(ns)=\frac{\alpha^5 m_\mu}{2n^3}.$$

The corrections to this expression are described by the diagrams in Fig. 7. The contribution corresponding to the radiative corrections of the bound muon–antimuon pair (Fig. 7a) is also similar to the correction for positronium:[25]

$$\Delta\Gamma_{\rm rad}(ns)=-\frac{\alpha}{\pi}\,\frac{20-\pi^2}{4}\Gamma^{(0)}_{\rm PM}(ns). \tag{32}$$

Another correction to two-photon decay corresponds to the Uehling correction to the wave function (Fig. 7b). It is state-dependent and proportional to (16) and (17). The results are listed in Table III. The final correction to the two-photon width is

$$\Delta\Gamma^{2\gamma}_{\rm PM}(1s)=\frac{\alpha}{\pi}(-1.47)\Gamma^{(0)}_{\rm PM}(1s) \tag{33}$$

and

$$\Delta\Gamma^{2\gamma}_{\rm PM}(2s)=\frac{\alpha}{\pi}(-1.61)\Gamma^{(0)}_{\rm PM}(2s)\,. \tag{34}$$

The third diagram in Fig. 7 (diagram c) corresponds to another decay channel: one of the photons is transformed into an electron–positron pair, and the final state turns out to be three-particle, $ee\gamma$. It is convenient to do the necessary calculations with the width interpreted as the imaginary part of the energy. Here the leading contribution corresponds to the imaginary parts of the diagrams in Fig. 8a, and the desired correction, to the imaginary parts of the diagrams in Fig. 8b. For the polarization insertion we again use the dispersion integral (9). The order in which the integral and imaginary part are calculated can be reversed. As a result we have

$$\Delta\Gamma^{ee\gamma}_{\rm PM}(ns)=\frac{\alpha}{\pi}\int_0^1 dv\,\frac{v^2(1-v^2/3)}{1-v^2}\Gamma^{(0)}(\lambda,0), \tag{35}$$

where $\Gamma^{(0)}(\lambda,0)$ is the width of paradimuonium decay into real and virtual photons with mass (8), or in the limit $m_\mu\gg m_e$,

$$\Delta\Gamma^{ee\gamma}_{\rm PM}(ns)=\frac{\alpha}{\pi}\left\{\left(\frac{2}{3}\ln 2-\frac{5}{9}\right)\Gamma^{(0)}(0,0)\right.$$

$$\left.+\frac{1}{3}\int_{4m_e^2}^\infty\frac{d\lambda^2}{\lambda^2}\Gamma^{(0)}(\lambda,0)\right\}. \tag{36}$$

Note that the real width of two-photon decay is

$$\Gamma^{(0)}_{\rm PM}=\tfrac{1}{2}\Gamma^{(0)}(0,0). \tag{37}$$

The combinatorial factor appears because the PM→2γ decay actually involves two identical massless photons,

TABLE III. Corrections of order $\alpha^6 m$ to the decay width of the $1s$ and $2s$ levels in paradimuonium are given in relative units: $\Delta\Gamma_{\rm PM}(ns)=(\alpha/\pi)C\Gamma^{(0)}_{\rm PM}(ns)$. The notation of the contributions corresponds to Fig. 7.

Contribution	$C(1s)$	$C(2s)$
a	−2.53	−2.53
b	1.06	0.92
c	6.26	6.26
Total	4.79	4.65

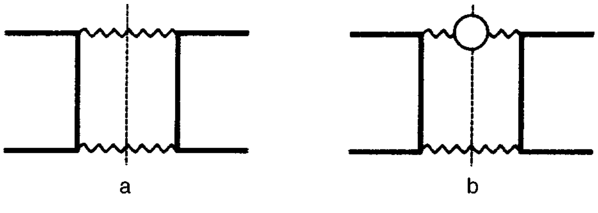


FIG. 8. Some diagrams for the width of paradimuonium decay as the imaginary part of the energy: (a) the leading contribution, and (b) decay with production of a pair and a photon.

TABLE IV. Properties of the low-lying states of dimuonium as compound particles. Decay modes with relative probability of at least 1% are included.

State	Mass, MeV	J^{PC}	τ, s	Decay channel	$\Delta\Gamma/\Gamma$, %
PM 1s	211.3154	0^{-+}	$0.59547(33)\times10^{-12}$	$\gamma\gamma$	98.5
				$ee\gamma$	1.5
2s	211.3164	0^{-+}	$4.7653(26)\times10^{-12}$	$\gamma\gamma$	98.5
				$ee\gamma$	1.5
OM 1s	211.3154	1^{--}	$1.79073(23)\times10^{-12}$	ee	86(4)
				$ee\gamma$	13(4)
				$ee\gamma\gamma$	~1
2s	211.3164	1^{--}	$14.3305(19)\times10^{-12}$	ee	86(4)
				$ee\gamma$	13(4)
				$ee\gamma\gamma$	~1

while in calculating $\Gamma^{(0)}(\lambda,0)$ in the limit $\lambda=0$, the photons are assumed to be distinct. Clearly, if we ignore the photon mass λ in $\Gamma^{(0)}(\lambda,0)$, the integral with respect to the dispersion variable diverges at large λ (i.e., $v\simeq1$). The divergence is cut off in $\Gamma^{(0)}(\lambda,0)$ kinematically; decay is possible only if $\lambda<2m_\mu$. As a result, instead of a divergence we have a logarithmic contribution to the width. The constant term is calculated in Appendix B. The result is

$$\Delta\Gamma^{ee\gamma}_{\mathrm{PM}}(ns)=\frac{\alpha}{\pi}\left(\frac{4}{3}\ln\frac{2m_\mu}{m_e}-\frac{16}{9}\right)\frac{1}{n^3}\Gamma^{(0)}_{\mathrm{PM}}(1s), \tag{38}$$

which amounts to 1.5% of the total width.

The final result for parapositronium is

$$\Delta\Gamma_{\mathrm{PM}}(1s)=\frac{\alpha}{\pi}\,4.79\Gamma^{(0)}_{\mathrm{PM}}(1s) \tag{39}$$

and

$$\Delta\Gamma_{\mathrm{PM}}(2s)=\frac{\alpha}{\pi}\,4.65\Gamma^{(0)}_{\mathrm{PM}}(1s). \tag{40}$$

5. DISCUSSION

First and foremost, we give the expressions for the lifetimes and hyperfine splitting of the lower levels with allowance for radiative corrections:

$$\tau_{\mathrm{OM}}(1s)=1.7907(8)\times10^{-12}\ \mathrm{s}, \tag{41}$$

$$\tau_{\mathrm{OM}}(2s)=14.331(6)\times10^{-12}\ \mathrm{s}, \tag{42}$$

$$\tau_{\mathrm{PM}}(1s)=0.59547(33)\times10^{-12}\ \mathrm{s}, \tag{43}$$

$$\tau_{\mathrm{PM}}(2s)=4.7653(26)\times10^{-12}\ \mathrm{s}. \tag{44}$$

The contributions of the individual terms are listed in Tables II and III. The contributions of the next order are estimated to be 5% of the one-loop contributions. The annihilation lifetimes of dimuonium states must be compared with the muon lifetime ($\tau_\mu=2.20\times10^{-6}$ s) and the atomic lifetimes of higher states. For instance, the annihilation width of the $2p$ level differs from the width of s states by an additional factor α^2, with the result that the lifetime of this level is completely determined by the one-photon transition[4] $2p\rightarrow1s$:

$$\tau(2p\rightarrow1s)=15.4\times10^{-12}\ \mathrm{s}, \tag{45}$$

and turns out to be of the same order as the annihilation lifetimes for the $1s$ and $2s$ levels given above.

Note that the sign of corrections of relative order α for the widths and for the hyperfine splitting is opposite that of the correction for positronium. In the case of hyperfine splitting (Table I), this is due to the large contribution of the insertion of electron vacuum polarization into an annihilation photon. As explained earlier, the logarithmic contribution is also present in the decay of paradimuonium, and it is related to the asymptotic behavior of the polarization operator.

We assume that dimuonium can be studied by the methods of experimental high-energy particle physics, and the properties of the dimuonium low-lying states as (compound) particles are listed in Table IV. For orthodimuonium we added the estimate for the decay channel OM→$ee\gamma\gamma$, which appears in the next perturbation order. Note that one of the decay channels, which emerges as a first-order correction in perturbation theory (OM→$ee\gamma$), amounts to 15%, and therefore can be used for detection purposes together with the principal channel.

Hyperfine splitting in dimuonium can also be measured by methods of elementary particle physics as the interference of the *ortho* and *para* states in the field of an IR laser. The results for hyperfine splitting are

$$E_{\mathrm{hfs}}(1s)=4.23284(35)\times10^{7}\ \mathrm{MHz}, \tag{46}$$

$$E_{\mathrm{hfs}}(2s)=5.28940(34)\times10^{6}\ \mathrm{MHz}. \tag{47}$$

Measurements of lifetimes and hyperfine splitting to 1% accuracy will make it possible to check the predictions of quantum electrodynamics. The corrections amount to about 1% for widths and roughly 0.1% for hyperfine splitting. In discussing the various QED contributions, one must bear in mind that some of the corrections for hyperfine splitting and *para* state decay are the same for dimuonium and positronium, while others yield contributions only for dimuonium. Positronium is a well-studied system (see Appendix C), with the result that measurements for dimuonium make it possible to check the QED contributions specific to dimuonium, which amount to about 0.7% for hyperfine splitting and 1.7% for the width of paradimuonium decay. To check the QED prediction one can measure partial widths, which for some (nonleading) channels may amount to 15% of the total width. Measurements of all these quantities will make it possible to test QED at large time-like momenta (in comparison to the electron mass) and to measure the hadronic vacuum polarization at $q^2=4m_\mu^2$, whose contribution amounts to 0.1% of the orthodimuonium decay width.

In conclusion we summarize the main properties of dimuonium that are of interest from an experimental standpoint. The spectrum of dimuonium states is a family of neutral particles that interfere in the field of an IR laser. The lifetimes lie in the picosecond range, and the decay products are ultrarelativistic electrons and positrons and hard gamma photons (see Table IV). Thus, the bound muon–antimuon system can serve as a good test particle for perfecting detection methods for exotic neutral elementary particles, which interact only weakly with normal matter.

Part of the present work was done during the stay of one of the authors (S.G.K.) in Dresden, Germany. He is grateful to the staffs of the Max Planck Institute of Physics of Complex Systems and the Dresden Technical University for their hospitality and support. This work was partially supported by the following grants: DFG SO333/1-2(U.D.J. and G.S.), and 95-02-03977 from the Russian Fund for Fundamental Research (S.G.K. and V.G.I).

6. ADDENDUM

After this paper was accepted for publication, the work of Owen and Repko[34] of 1972 was brought to our attention, for which we are grateful to W. W. Repko. The estimate obtained in that work for the contribution of the insertion of the hadronic vacuum polarization to hyperfine splitting is consistent with our more accurate calculations. We would also like to note the papers by Holvik and Olsen and Lyuboshits[35] in addition to those cited in the Introduction in connection with dimuonium formation.[8,10]

APPENDIX A

Contribution of hadronic polarization to hyperfine splitting

The contribution of hadronic polarization to hyperfine splitting and the width of orthodimuonium decay can be conveniently written as follows:

$$\Delta E_{\mathrm{hadr}}(ns)=\frac{\alpha}{\pi}C\,\frac{E_F}{n^3}=\frac{3}{7}\,\frac{\alpha}{\pi}C'_{\mathrm{hadr}}\frac{E_F}{n^3}, \tag{48}$$

$$\Delta\Gamma_{\mathrm{hadr}}(ns)=2\,\frac{\alpha}{\pi}C'_{\mathrm{hadr}}\Gamma^{(0)}_{\mathrm{OM}}(ns). \tag{49}$$

The coefficient C'_{hadr} is the ratio of the free electron propagator to the correction with the insertion of vacuum polarization for timelike momentum $q^2=(2m_\mu)^2$.

The correction to the propagator can be represented via the dispersion integral (7). Following Ref. 20, we separate out the π-meson contribution from the spectral function $\rho(s)$. The contribution is

$$\rho_{\pi\pi}(s)=\frac{(s-4m_\pi^2)^{3/2}}{12s^{5/2}}|F_\pi(s)|^2, \tag{50}$$

where $F_\pi(s)$ is the π-meson form factor, which is required because of the strong interaction inside the pion loop. The form factor is used in the parametrization scheme of Gounaris and Sakurai:[21]

$$F_\pi(s)=(m_\rho^2+dm_\rho\Gamma_\rho)\left\{(m_\rho^2-s)+\left(\Gamma_\rho\frac{m_\rho^2}{k_\rho^3}[k(s)^2(h(s)\right.\right.$$
$$\left.-h_\rho)+k_\rho^2h'(m_\rho^2)(m_\rho^2-s)]\right)$$
$$\left.-i\left[m_\rho\Gamma_\rho\left(\frac{k(s)}{k_\rho}\right)^3\frac{m_\rho}{\sqrt{s}}\right]\right\}^{-1}, \tag{51}$$

where we have introduced the auxiliary quantities

$$d=\frac{3}{\pi}\,\frac{m_\pi^2}{k_\rho^2}\ln\left(\frac{m_\rho+2k_\rho}{2m_\pi}\right)+\frac{m_\rho}{2\pi k_\rho}-\frac{m_\pi^2m_\rho}{\pi k_\rho^3}\approx 0.48, \tag{52}$$

$$k(s)=\frac{1}{2}\sqrt{s-4m_\pi^2}, \tag{53}$$

$$h(s)=\frac{2}{\pi}\,\frac{k(s)}{\sqrt{s}}\ln\frac{\sqrt{s}+2k(s)}{2m_\pi}, \tag{54}$$

$h'(s)$ is the s-derivative of $h(s)$, $k_\rho=k(m_\rho^2)$, and $h_\rho=h(m_\rho^2)$. In calculating the contribution of the pion loop we took $\Gamma_\rho=150.7(1.2)$ MeV and $m_\rho=768.5(6)$ MeV (see Ref. 7). The result is

$$C'_{\pi\pi}=4m_\mu^2\int_{4m_\pi^2}^{\infty}ds\,\frac{\rho_{\pi\pi}(s)}{4m_\mu^2-s}=-0.128. \tag{55}$$

Clearly, the size of the contribution is determined by the ρ-meson pole. Indeed, in the pole approximation,

$$\rho_\rho(s)=\frac{4\pi^2}{f_\rho^2}\,\delta(s-m_\rho^2),$$

and the result at $f_\rho^2/4\pi=2.2$ (see Ref. 22),

$$C'_\rho=-0.116, \tag{56}$$

is close to the one obtained earlier.

Next, to (55) we must add the contribution of the other resonances (ω and ϕ). This can be done in the pole approximation. The corresponding values of the parameters are $f_\omega^2/4\pi=18(2)$ and $f_\phi^2/4\pi=11(2)$ (Refs. 20, and 22), $m_\omega=782$ MeV and $m_\phi=1019$ MeV (Ref. 7). The results are listed in Table V.

Now we must add the nonresonant contribution. We estimate the contribution of the background in the following way: the weighting function is defined to be

$$\rho_>(s)=\frac{R(s)}{3s}, \tag{57}$$

where

$$R=\frac{\sigma(e^+e^-\to\mathrm{hadrons})}{\sigma(e^+e^-\to\mu^+\mu^-)}, \tag{58}$$

TABLE V. Contributions of hadronic vacuum polarization to the hyperfine splitting and width of orthodimuonium: $\pi\pi$, the contribution of the pionic loop with allowance for interaction; ω and ϕ, contributions of the corresponding resonances in the pole approximation; the background corresponds to nonresonant contributions to the dispersion integral above 1 GeV. The corrections are given in terms of C'_{hadr} (see Eqs. (48) and (49)).

Contribution	C'
$\pi\pi$	$-0.128(13)$
ω	$-0.014(3)$
ϕ	$-0.012(2)$
Background (1–4 GeV)	$-0.028(6)$
Background (above 4 GeV)	0.04
Total	$-0.186(22)$

and the value of the auxiliary function $R(s)$ is taken approximately equal to 2 for $\sqrt{s}$ between 1 and 4 GeV, and to 4 for $\sqrt{s}$ larger than 4 GeV (see, e.g., Ref. 7).

All the results for hadronic polarization are listed in Table V. They amount to

$$\Delta E_{\mathrm{hadr}}(ns)=\frac{\alpha}{\pi}[-0.080(9)]\,\frac{E_F}{n^3}, \tag{59}$$

or 0.02% for hyperfine splitting, and

$$\Delta\Gamma_{\mathrm{hadr}}(ns)=\frac{\alpha}{\pi}[-0.37(4)]\Gamma^{(0)}_{\mathrm{OM}}(ns), \tag{60}$$

or 0.09% for the orthopositronium decay width.

APPENDIX B

Decay of paradimuonium into a photon and an electron-positron pair

As noted earlier, to calculate the probability of the PM $\to ee\gamma$ decay we must evaluate the integral (36). Consider the quantity

$$C_{ee\gamma}=\frac{\alpha}{\pi}\left\{\left(\frac{2}{3}\ln 2-\frac{5}{9}\right)\frac{\Gamma^{(0)}(0,0)}{\Gamma^{(0)}_{\mathrm{PM}}}\right.$$
$$\left.+\frac{1}{3}\int_{4m_e^2}^{\infty}\frac{d\lambda^2}{\lambda^2}\,\frac{\Gamma^{(0)}(\lambda,0)}{\Gamma^{(0)}_{\mathrm{PM}}}\right\}, \tag{61}$$

or

$$C_{ee\gamma}=2\,\frac{\alpha}{\pi}\left\{\left(\frac{2}{3}\ln 2-\frac{5}{9}\right)+\frac{1}{3}\int_{4m_e^2}^{\infty}\frac{d\lambda^2}{\lambda^2}\,\frac{\Gamma^{(0)}(\lambda,0)}{\Gamma^{(0)}(0,0)}\right\}. \tag{62}$$

Let us discuss the amplitude of emission of two photons, with one photon being massless (we denote its 4-momentum by k_ν). In the atom's center-of-mass frame, we can ignore the three-dimensional momenta of the muon and antimuon, and following the standard procedure we can easily see that the amplitude is proportional to

$$\frac{\gamma_i(\gamma_j k_j)\gamma_l}{2k_0 m},$$

where the indices of all γ-matrices are purely spatial. Since the photon is real, this quantity does not depend on the magnitude of the momentum $\mathbf{k}$, just on its direction. The dependence on the "mass" λ of one of the photons is now contained only in the photon propagators and integrals with respect to energy and the magnitude of the momentum, which can be "removed" by delta functions when we pass to the imaginary part:

$$X(\lambda)=\int k_0\int d|\mathbf{k}|\;\mathbf{k}^2\delta(k^2)\,\delta(k_1^2)\,\theta(k_0)\,\theta(2m_\mu-k_0),$$

where $k_1=(2m_\mu-k_0,-\mathbf{k})$ is the 4-momentum of the massive photon, with

$$k_1^2=4m_\mu^2-4m_\mu k_0+k^2.$$

The kinematics of emission is entirely determined by the delta functions and corresponds to the condition

$$|\mathbf{k}|=m_\mu-\frac{\lambda^2}{4m_\mu},$$

which leads to

$$X(\lambda)=\frac{1}{8}\left(1-\frac{\lambda^2}{4m_\mu^2}\right)\theta(2m_\mu-\lambda).$$

Note that there are no other integrations: the kinematics fixes everything except the direction of emission, but since the photons emerge in opposite directions, there is no relative angle, and after averaging over the polarizations of the gamma photons, summation over the angles is trivial. As a result we arrive at the simple integral

$$C_{ee\gamma}=2\,\frac{\alpha}{\pi}\left\{\left(\frac{2}{3}\ln 2-\frac{5}{9}\right)+\frac{1}{3}\int_{4m_e^2}^{\infty}\frac{d\lambda^2}{\lambda^2}\,\frac{X(\lambda)}{X(0)}\right\}$$
$$=2\,\frac{\alpha}{\pi}\left\{\left(\frac{2}{3}\ln 2-\frac{5}{9}\right)+\frac{1}{3}\int_{4m_e^2}^{4m_\mu^2}\frac{d\lambda^2}{\lambda^2}\left(1-\frac{\lambda^2}{4m_\mu^2}\right)\right\}, \tag{63}$$

and calculating the asymptotic behavior is easy and leads to (38).

APPENDIX C

Hyperfine splitting and the decay of positronium

Here we discuss the current status of theory and experiment for the hyperfine splitting of the ground state of positronium and for parapositronium decay. The decay of parapositronium in leading approximation differs from that of dimuonium, so comparison with the decay of dimuonium is of no interest.

C.1. Hyperfine splitting of the ground state in positronium

The hyperfine splitting of the ground state in positronium is described by the following expression:[27,28]

$$E^{\mathrm{Ps}}_{\mathrm{hfs}}(1s)=\left(1-2.1179\,\frac{\alpha}{\pi}+0.357\alpha^2\,\ln\frac{1}{\alpha}\right.$$
$$\left.+[-0.84(6)+C_2]-1.5\alpha^3\,\ln^2\frac{1}{\alpha}\right)E_F, \tag{64}$$

where the Fermi energy was defined earlier in the paper, and the unknown coefficient C_2 corresponds to a two-loop correction to one-photon annihilation and the contribution associated with that correction. Numerically, the theoretical expression leads to the result (see Ref. 27)[5]

$$E^{\rm Ps}_{\rm hfs}(1s)=[203395.3(7)+4.0\ C_2]\ \ {\rm MHz}. \tag{65}$$

This should be compared with the experimental results, which are 203387.0(16) MHz,[29] 203384.9(12) MHz,[30] and 203389.1(7) MHz.[31] The estimated experimental error is $\alpha^2 E_F$. At present, calculations of the coefficient C_2 are being completed.[6]

C.2. Decay of parapositronium

The accuracy of calculating and measuring the width of parapositronium decay is not so high. The theoretical expression is[32]

$$\Gamma_{\rm PP}(1s)=\Gamma^{(0)}_{\rm PP}(1s)\left(1-2.533\,\frac{\alpha}{\pi}+2\alpha^2\,\ln\frac{1}{\alpha}\right)\Gamma^{(0)}_{\rm PP}(1s), \tag{66}$$

where the leading term is defined by analogy with (6). Corrections of order α^2 are not known. The experimental result[33] has an error of 2×10^{-4} in relative units, and is consistent with the theoretical result.

[*] E-mail: sgk@onti.vniim.spb.su

[1] We call this system dimuonium, distinguishing between states with zero total spin $S=0$ (paradimuonium, or PM) and unit total spin $S=1$ (orthodimuonium, or OM). The name *muonium* is reserved for an atom consisting of an electron and antimuon.

[2] We use the relativistic system of units, in which $\hbar=c=1$ and $\alpha=e^2$.

[3] Bearing in mind that $E=h\nu$, we present the results for hyperfine splitting in terms of ν.

[4] Here we have allowed for the dependence of the radiative width on the mass of the nucleus, which in the nonrelativistic approximation leads to the result $(1+Zm/M)^2(m_r^3/m^3)\Gamma_\infty$, where Γ_∞ is the probability for a nucleus with infinite mass.[26]

[5] G. Adkins and P. Labelle (private communication).

[6] G. Adkins (private communication).

---

[1] R. Coombes, R. Flexer, A. Hall *et al.*, Phys. Rev. Lett. **37**, 249 (1976).

[2] S. H. Aronson, R. H. Bernstein, G. J. Bock *et al.*, Phys. Rev. Lett. **48**, 1078 (1982); Phys. Rev. D **33**, 3180 (1986).

[3] G. D. Alekseev, A. I. Baranovskiĭ, O. E. Gorchakov, K. I. Gubrienko, V. V. Karpukhin, V. I. Kotov, V. V. Kruglov, A. V. Kulikov, A. B. Kuptsov, L. L. Nemenov, V. M. Plotko, E. E. Polbennikov, S. V. Trusov, D. M. Khazins, A. S. Chvyrov, M. N. Shumakov, and S. M. Frolikov, Yad. Fiz. **40**, 139 (1984) [Sov. J. Nucl. Phys. **40**, 87 (1984)]; L. G. Afanas'ev, N. I. Balalykin, O. E. Gorchakov, V. V. Karpukhin, V. I. Komarov, A. V. Kolomyĭchenko, V. I. Kotov, V. V. Kruglov, A. V. Kulikov, A. V. Kuptsov, L. L. Nemenov, M. V. Nikitin, Zh. P. Pustyl'nik, A. V. Skrypnik, S. V. Trusov, A. S. Chvyrov, M. N. Shumakov, and S. M. Frolikov, Yad. Fiz. **50**, 7 (1989) [Sov. J. Nucl. Phys. **50**, 4 (1989)].

[4] L. G. Afanasyev, A. S. Chvyrov, V. V. Karpukhin *et al.*, Phys. Lett. B **236**, 116 (1990); L. G. Afanas'ev, O. E. Gorchakov, V. V. Karpukhin, V. I. Komarov, A. V. Kolomyĭchenko, V. V. Kruglov, A. V. Kulikov, A. V. Kuptsov, L. L. Nemenov, M. V. Nikitin, Zh. P. Pustyl'nik, S. V. Trusov, M. N. Shumakov, S. M. Frolikov, and A. S. Chvyrov, Yad. Fiz. **51**, 1040 (1990) [Sov. J. Nucl. Phys. **51**, 664 (1990)].

[5] L. G. Afanasyev, A. S. Chvyrov, O. E. Gorchakov *et al.*, Phys. Lett. B **308**, 200 (1993); L. G. Afanasyev, A. S. Chvyrov, O. E. Gorchakov *et al.*, Phys. Lett. B **338**, 478 (1994).

[6] L. R. Wiencke, M. D. Church, E. E. Gottschalk *et al.*, Phys. Rev. D **46**, 3709 (1992); J. Uribe, E. P. Hartouni, D. A. Jensen *et al.*, Phys. Rev. D **49**, 4373 (1994); L. R. Wiencke, Ph.D. Thesis, Columbia University (1993).

[7] R. M. Barnett, C. D. Carone, D. E. Groom, T. G. Trippe, C. G. Wohl, B. Armstrong, P. S. Gee, G. S. Wagman, F. James, M. Mangano, K. Mönig, L. Montanet, J. L. Feng, H. Murayama, J. J. Hernández, A. Manohar, M. Aguilar-Benitez, C. Caso, R. L. Crawford, M. Roos, N. A. Törnqvist, K. G. Hayes, K. Hagiwara, K. Nakamura, M. Tanabashi, K. Olive, K. Honscheid, P. R. Burchat, R. E. Shrock, S. Eidelman, R. H. Schindler, A. Gurtu, K. Hikasa, G. Conforto, R. L. Workman, C. Grab, and C. Amsler, Phys. Rev. D **54**, 1 (1996).

[8] S. M. Bilen'kiĭ, N. V. Kh'eu, L. L. Nemenov, and F. G. Tkebuchava, Yad. Fiz. **10**, 812 (1969).

[9] S. Wycech and A. M. Green, Nucl. Phys. A **562**, 446 (1993); Z. K. Silagadze, JETP Lett. **60**, 689 (1994); V. Lyubovitskij and A. Rusetsky, Phys. Lett. B **389**, 181 (1996); O. E. Gorchakov, A. V. Kuptsov, L. L. Nemenov, and D. Yu. Ryabikov, Yad. Fiz. **59**, 2015 (1996) [Phys. At. Nucl. **59**, 1942 (1996)].

[10] L. L. Nemenov, Yad. Fiz. **15**, 1047 (1972) [Sov. J. Nucl. Phys. **15**, 582 (1972)]; G. A. Kozlov, Yad. Fiz. **48**, 265 (1988) [Sov. J. Nucl. Phys. **48**, 167 (1988)].

[11] L. L. Nemenov, Yad. Fiz. 16, 125 (1972) [Sov. J. Nucl. Phys. **16**, 67 (1973)]; Yad. Fiz. **41**, 981 (1985) [Sov. J. Nucl. Phys. **41**, 629 (1985)].

[12] J. Malefant, Phys. Rev. D **36**, 863 (1987).

[13] R. Karplus and A. Klein, Phys. Rev. **87**, 848 (1952).

[14] T. Fulton and P. C. Martin, Phys. Rev. **95**, 811 (1954).

[15] J. A. Wheeler, Ann. (N.Y.) Acad. Sci. **95**, 219 (1946).

[16] S. G. Karshenboĭm, Yad. Fiz. 56, 115 (1993) [Phys. At. Nucl. **56**, 777 (1993)].

[17] V. G. Ivanov and S. G. Karshenboĭm, Zh. Éksp. Teor. Fiz. **112**, 805 (1997) [JETP **85**, 435 (1997)].

[18] V. G. Ivanov and S. G. Karshenboĭm, Zh. Éksp. Teor. Fiz. **109**, 1219 (1996) [JETP **82**, 656 (1996)]; V. G. Ivanov and S. G. Karshenboim, Phys. Lett. A **210**, 313 (1996).

[19] A. Ore and J. Powell, Phys. Rev. **75**, 1696 (1949).

[20] J. R. Sapirstein, E. A. Terray, and D. R. Yennie, Phys. Rev. D **29**, 2290 (1984).

[21] G. I. Gounaris and J. J. Sakurai, Phys. Rev. Lett. **21**, 244 (1968).

[22] T. H. Bauer, R. D. Spital, D. R. Yennie, and F. M. Pipkin, Rev. Mod. Phys. **50**, 261 (1978).

[23] J. Schwinger, *Particles, Sources and Fields*, Addison-Wesley, Reading, Mass. (1970).

[24] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 2nd ed., Pergamon Press, Oxford (1982).

[25] I. Harris and L. M. Brown, Phys. Rev. **105**, 1656 (1957).

[26] Z. Fried and A. D. Martin, Nuovo Cimento **29**, 574 (1963).

[27] S. G. Karshenboĭm, Yad. Fiz. **56**, 155 (1993).

[28] S. G. Karshenboĭm, Zh. Éksp. Teor. Fiz. **103**, 1105 (1993) [JETP **76**, 541 (1993)]; G. S. Adkins, Y. M. Aksu, and M. H. T. Bui, Phys. Rev. A **47**, 2640 (1993).

[29] D. L. Mills and G. H. Bearman, Phys. Rev. Lett. **34**, 246 (1975).

[30] P. O. Egan, V. W. Hughes, and M. H. Yam, Phys. Rev. A **15**, 251 (1977).

[31] M. Ritter, P. O. Egan, V. W. Hughes, and K. A. Woodle, Phys. Rev. A **30**, 1331 (1984).

[32] I. B. Khriplovich and A. S. Yelkhovsky, Phys. Lett. B **246**, 520 (1990).

[33] A. H. Al-Ramadhan and D. W. Gidley, Phys. Rev. Lett. **72**, 1632 (1994).

[34] D. A. Owen and W. W. Repko, Phys. Rev. A **5**, 1570 (1972).

[35] E. Holvik and H. A. Olsen, Phys. Rev. D **35**, 2124 (1987); V. L. Lyuboshits, Yad. Fiz. **45**, 1099 (1987) [Sov. J. Nucl. Phys. **45**, 682 (1987)].

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Theory of transfer with delay for trapping of nonstationary acoustic radiation in a resonant randomly inhomogeneous medium

Yu. N. Barabanenkov

''Éldis'' Scientific-Research Center of Electronic Diagnostics Systems, Russian Academy of Sciences, 101000 Moscow, Russia

M. Yu. Barabanenkov

Institute of Problems of Microelectronics and Ultrapure Materials Technology, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 17 January 1997)
Zh. Éksp. Teor. Fiz. **113**, 432–444 (February 1998)

The propagation of a quasimonochromatic wave packet of acoustic radiation in a discrete randomly-inhomogeneous medium under the condition that the carrier frequency of the packet is close to the resonance frequency of Mie scattering by an isolated scatterer is studied. The two-frequency Bethe–Salpeter equation in the form of an exact kinetic equation that takes account of the accumulation of the acoustic energy of the radiation inside the scatterers is taken as the initial equation. This kinetic equation is simplified by using the model of resonant point scatterers, the approximation of low scatterer density, and the Fraunhofer approximation in the theory of multiple scattering of waves. This leads to a new transport equation for nonstationary radiation with three Lorentzian delay kernels. In contrast to the well-known Sobolev radiative transfer equation with one Lorentzian delay kernel, the new transfer equation takes account of the accumulation of radiation energy inside the scatterers and is consistent with the Poynting theorem for nonstationary acoustic radiation. The transfer equation obtained with three Lorentzian delay kernels is used to study the Compton–Milne effect—trapping of a pulse of acoustic radiation diffusely reflected from a semi-infinite resonant randomly-inhomogeneous medium, when the pulse can spend most of its propagation time in the medium being ''trapped'' inside the scatterers. This specific albedo problem for the transfer equation obtained is solved by applying a generalized nonstationary invariance principle. As a result, the function describing the scattering of a diffusely reflected pulse can be expressed in terms of a generalized nonstationary Chandrasekhar H-function, satisfying a nonlinear integral equation. Simple analytical asymptotic expressions are found for the scattering function for the leading and trailing edges of a diffusely reflected δ-pulse as functions of time, the reflection angle, the mean scattering time of the radiation, the elementary delay time, and the parameter describing the accumulation of radiation energy inside the scatterers. These asymptotic expressions demonstrate quantitatively the retardation of the growth of the leading edge and the retardation of the decay of the trailing edge of a diffusely reflected δ-pulse when the conventional radiative transfer regime goes over to a regime of radiation trapping in a resonant randomly-inhomogeneous medium. [S1063-7761(98)00402-8]

1. INTRODUCTION

Over the past decade the multiple scattering of classical wave fields in discrete randomly inhomogeneous media has attracted great interest from theoreticians and experimenters because of the possibility of Anderson localization of such wave fields, specifically, acoustic waves.[1,2] In searching for localization of an acoustic wave packet in the case of three dimensions it is desirable to employ a medium consisting of an ensemble of spherical scatterers randomly distributed in space under the condition that the carrier frequency of the packet is close to the resonant frequency for Mie scattering by an isolated scatterer. One would think that this resonance condition together with a sufficiently high spatial density of the scatterers could give, according to Refs. 1 and 3, a minimum transport mean free path length l_{tr} for the radiation and, correspondingly, a low value of the radiation diffusion coefficient D. However, resonance scattering is characterized not only by a large scattering cross section but also by a long radiation delay time.[4] For this reason additional difficulties can arise in determining the closeness to the localization threshold in a specific experiment, as noted in Ref. 5.

The problem is that according to experiments[6] the radiation diffusion coefficient in this resonant randomly inhomogeneous medium has the form $D=v_E l_{tr}/3$, where, on account of the above-mentioned time delay in an elementary resonance scattering event, the acoustic energy transport velocity v_E can be an order of magnitude lower than the phase velocity C_0 of sound in a homogeneous medium with no scatterers. Such a decrease in the energy transport velocity of classical, specifically, acoustic, wave fields with multiple resonance Mie scattering by an ensemble of randomly distributed scatterers, which has no relation to the phenomenon of wave localization, is physically closely related, as can be easily seen from Ref. 7, to the well-known Compton[8] and Milne[9] trapping of optical resonance radiation in gases whose atoms possess a resonance absorption line.

A definition of trapping of resonance radiation in gases was given by Compton.[8] If the radiation delay time t_{del} in an elementary resonance scattering event is longer than the free flight time t_l of the radiation between scattering events,

$$t_{\mathrm{del}}/t_l \gg 1, \tag{1}$$

then the radiation propagating in the medium spends most of its time ''trapped'' inside the resonant scatterers. It should be noted that most treatments[8–14] on the theory of the trapping effect focus mainly on the transfer of an excited state of gas atoms by means of trapped resonance radiation. A substantial advance in extending the conventional theory of radiation transfer to this effect was made by Sobolev[15] and his successors.[16–18] On the basis of phenomenological considerations Sobolev derived a radiative transfer equation with an exponential (Lorentzian) delay kernel in a term with an isotropic scattering phase function for the radiant intensity of nonstationary radiation. Although Sobolev's theory is physically clear and mathematically well developed,[15–21] it is not completely systematic and needs to be improved. Indeed, in Sobolev's radiative transfer equation the delay effect is taken into account only in the term with the scattering phase function and is neglected in the term with the extinction coefficient. The accumulation of radiation energy inside resonance scatterers, which was studied in Refs. 5, 6, 22, and 23, is not discussed at all in Sobolev's theory. As a result of these deficiencies, Sobolev's radiative transfer equation with delay is inconsistent with Poynting's theorem for nonstationary radiation.

Our objective in the present paper is to eliminate these deficiencies of Sobolev's theory of radiative transfer with delay. The deficiences of Sobolev's theory for the trapping of resonance radiation are due to the phenomenological character of the theory and can be eliminated by returning to the ''first principles'' of the statistical theory of multiple scattering of waves in randomly inhomogeneous media. This is demonstrated in the present paper for the case of resonant multiple scattering of a quasimonochromatic acoustic wave packet by an ensemble of randomly distributed scatterers.

The starting point is the exact generalized kinetic equation,[24] which follows from the two-frequency Bethe-Salpeter equation for the space–time coherence spectral density function of an acoustic field. The generalized two-frequency Ward–Takahashi identity[24–26] for a classical wave field in a randomly-inhomogeneous medium, which obeys a hyperbolic stochastic equation, permits introducing into the kinetic equation an operator that takes account of the accumulation of acoustic energy inside the scatterers (Sec. 2). A modified radiation transfer equation with three Lorentzian delay kernels is derived by applying to the kinetic equation the low-density approximation in the model of resonant point scatterers[27] together with the Fraunhofer approximation[28,29] (Sec. 3). The albedo problem of diffuse reflection of an acoustic pulse from a resonant semi-infinite medium is solved for the modified equation obtained (Sec. 4).

2. GENERALIZED KINETIC EQUATION FOR MULTIPLE SCATTERING OF ACOUSTIC WAVES

Our starting point is the stationary wave equation for the pressure $P(\mathbf{r},t)$ in a sound wave propagating in a discrete randomly-inhomogeneous medium with phase velocity $C(\mathbf{r})$. Assuming that the density ρ_0 of the medium, unperturbed by the sound wave, is the same and uniform inside and outside the scatterers, the wave equation can be written in the form,[30]

$$-\Delta P+\frac{1}{C^2(\mathbf{r})}\frac{\partial^2 P}{\partial t^2}=-J, \tag{2}$$

where $J(\mathbf{r},t)$ is a fixed source of the acoustic disturbance. By analogy to the case of electromagnetic waves,[23] it is convenient to set $C_0^2/C^2(\mathbf{r})=\epsilon(\mathbf{r})$. Here the effective permittivity of the medium $\epsilon(\mathbf{r})=1+\delta\epsilon(\mathbf{r})$ includes a randomly-inhomogeneous component $\delta\epsilon(\mathbf{r})$

$$\delta\epsilon(\mathbf{r})=(\epsilon_1-1)\sum_i \chi(\mathbf{r}-\mathbf{r}_i). \tag{3}$$

Here ϵ_1 is the fixed permittivity of a scatterer and $\chi(\mathbf{r}-\mathbf{r}_i)$ is a characteristic function of the region of space occupied by the i-th scatterer centered at the point $\mathbf{r}_i$ and equals 1 inside and zero outside a scatterer. It is assumed that the ensemble of scatterers consists of identical spheres with fixed radius r_0 and that the centers $\mathbf{r}_i$ of the scatterers are randomly distributed in space.

Multiple scattering of nonstationary acoustic wave radiation in the discrete randomly inhomogeneous medium under study is described by the two-frequency Bethe–Salpeter equation.[24] As in to the case of electromagnetic radiation,[23] this equation can be transformed to the following generalized kinetic equation

$$\begin{aligned}\left\{-i\,\frac{\Omega\omega}{C_0^2}\,[1-A^+(\mathbf{p};\mathbf{q},\omega)]+i\mathbf{p}\cdot\mathbf{q}\right\}f(\mathbf{p};\mathbf{q},\omega)\\ =\int_{\mathbf{p}'}\Delta G(\mathbf{p}';\mathbf{q},\omega)K(\mathbf{p}',\mathbf{p};\mathbf{q},\omega)f(\mathbf{p};\mathbf{q},\omega)\\ -\Delta G(\mathbf{p};\mathbf{q},\omega)\int_{\mathbf{p}'}K(\mathbf{p},\mathbf{p}';\mathbf{q},\omega)f(\mathbf{p}';\mathbf{q},\omega)\\ -\Delta G(\mathbf{p};\mathbf{q},\omega)J(\mathbf{p};\mathbf{q},\omega),\end{aligned} \tag{4}$$

which is written in the spatial Fourier transform and temporal Laplace transform representations and the Wigner variables. The desired function $f(\mathbf{p};\mathbf{q},\omega)$ is a spatiotemporal coherence spectral density function of the sound field, $\langle P(\mathbf{R}+\mathbf{r}/2,T+t/2)P(\mathbf{R}-\mathbf{r}/2,T-t/2)\rangle$, where the angular brackets denote averaging over the ensemble, the wave vectors $\mathbf{p}$ and $\mathbf{q}$ correspond to the spatial vectors $\mathbf{r}$ and $\mathbf{R}$, and the frequencies Ω and ω correspond to the time variables t and T. The quantity $\Delta G(\mathbf{p};\mathbf{q},\omega)$ on the right-hand side of

Eq. (4) can be expressed in terms of the Fourier transform $G(\mathbf{p},\Omega)$ of the average Green's function by the relation

$$\Delta G(\mathbf{p};\mathbf{q},\omega)=\frac{1}{2i}\left[G_1(\mathbf{p}_+)-G_2(\mathbf{p}_-)\right]; \tag{5}$$

here (and below) the indices 1 and 2 correspond to the frequencies $\Omega_{1,2}=\pm\Omega+\omega/2+i0$ and the wave vectors $\mathbf{p}_\pm=\mathbf{p}\pm\mathbf{q}/2$. The intensity operator is denoted by $K(\mathbf{p},\mathbf{p}';\mathbf{q},\omega)$ and we have $J(\mathbf{p};\mathbf{q},\omega)=J_1(\mathbf{p}_+)J_2(-\mathbf{p}_-)$, where $J_{1,2}(\mathbf{p})$ are the Fourier transforms of $J_{1,2}(\mathbf{r})$. The symbol $\int_{\mathbf{p}}=(2\pi)^{-3}\int d\mathbf{p}$.

The most interesting quantity $A^+(\mathbf{p};\mathbf{q},\omega)$ in Eq. (4) is the Fourier transform of the operator $A^+(\mathbf{R};\mathbf{R}'+\mathbf{r}'/2,\mathbf{R}'-\mathbf{r}'/2;\Omega_1,\Omega_2)$ describing the accumulation of sound energy inside the scatterers. The wave vectors $\mathbf{p}$ and $\mathbf{q}$ correspond to the spatial vectors $\mathbf{r}'$ and $\mathbf{R}-\mathbf{R}'$. The general definition of the operator A^+ in the coordinate representation has the form[23]

$$A^+=\frac{1}{g_1+g_2}A_{12}^+, \tag{6}$$

where the operator A_{12}^+ and the analogous operator A_{12}^- are defined as

$$A_{12}^{\pm}=M_1\hat{\otimes}I\pm I\hat{\otimes}M_2\pm(G_1\hat{\otimes}I\pm I\hat{\otimes}G_2)K_{12}. \tag{7}$$

Here M is the mass operator, the function $g(\Omega)=\Omega^2/C_0^2$, I is the identity operator, and the symbol $\hat{\otimes}$ denotes a tensor product of two operators with the same first two arguments (see Ref. 23). The physical meaning of the operator (6) is clarified by the energy equation

$$\langle w(\mathbf{R},\omega)\rangle=\frac{1}{2\rho_0C_0^2}\int_{\mathbf{p}}\int_{\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{R}}[1-A^+(\mathbf{p};\mathbf{q},\omega)]f(\mathbf{p};\mathbf{q},\omega), \tag{8}$$

where $w=P_1P_2/(2\rho_0C_0^2)$ is the two-frequency spectral component of the sound energy density. The second term on the right-hand side of Eq. (8) shows that the operator (6) characterizes the spectral component of the sound energy accumulated inside the scatterers. The two other energy equations, related with the kinetic equation (4), have the form

$$\langle\mathbf{S}(\mathbf{R},\omega)\rangle=\frac{1}{2\rho_0\Omega}\int_{\mathbf{p}}\int_{\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{R}}\mathbf{p}f(\mathbf{p};\mathbf{q},\omega), \tag{9}$$

$$\langle Q(\mathbf{R},\omega)\rangle=\frac{1}{2\rho_0\Omega}\int_{\mathbf{p}}\int_{\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{R}}\Delta G(\mathbf{p};\mathbf{q},\omega)J(\mathbf{p};\mathbf{q},\omega). \tag{10}$$

On the left-hand sides of these equations $\mathbf{S}=(-P_1\nabla P_2+P_2\nabla P_1)/4i\rho_0\Omega$ and $Q=(P_1J_2-P_2J_1)/4i\rho_0\Omega$ are, respectively, the spectral components of the Poynting vector and of the intensity of the sources of the sound disturbance, taken with a minus sign. The energy equations (8)–(10) make it possible to verify that multiplying the kinetic equation (4) by $\exp(i\mathbf{q}\cdot\mathbf{R})$ and then integrating over the variables $\mathbf{p}$ and $\mathbf{q}$ gives an ensemble-averaged Poynting theorem for the initial wave equation (2) in the form

$$-i\omega\langle w\rangle+\mathrm{div}\langle\mathbf{S}\rangle+\langle Q\rangle=0. \tag{11}$$

The compatibility of the kinetic equation (4) with the averaged Poynting theorem (11) is based on the generalized two–frequency Ward–Takahashi identity

$$A_{12}^-=\frac{g_1-g_2}{g_1+g_2}A_{12}^+, \tag{12}$$

proved in Refs. 24–26.

3. MODIFIED ACOUSTIC RADIATIVE TRANSFER EQUATION WITH THREE LORENTZIAN DELAY KERNELS

We now return to the exact kinetic equation (4) in order to obtain from it an approximate modified radiation transport equation taking account of the delay effect. First we shall list all approximations that make it possible to switch from the kinetic equation (4) to the general modified radiative transfer equation with delay. Next, using in addition the low-density approximation and the model of resonant point scatterers,[27] we write out the modified equation with three Lorentzian delay kernels.

Spatial dispersion approximation

Let a quasimonochromatic wave packet (pulse) of acoustic radiation with spatial and temporal scales Δr_0 and Δt_0, where $k_0\Delta r_0\gg1$ and $\Omega\Delta t_0\gg1$, propagate in a discrete randomly inhomogeneous medium. The quantity $k_0=\Omega/C_0$ is the wave number in a homogeneous medium. Assuming that the intensity operator in Eq. (4) varies as a function of $\mathbf{q}$ on the scale $\Delta q\sim1/r_0$, satisfying the condition

$$r_0\ll\Delta r_0, \tag{13}$$

we can set $\mathbf{q}=0$ in all coefficients of the kinetic equation (4), except, of course, the coefficient $i\mathbf{p}\cdot\mathbf{q}$.

Strong delay approximation

The ratio of the time scale Δt_0 of the pulse and the scale $\Delta\omega\sim1/t_{\mathrm{del}}$ of the variation of the intensity operator in Eq. (4) as a function of ω is assumed to be arbitrary, i.e.,

$$0<t_{\mathrm{del}}/\Delta t_0<\infty. \tag{14}$$

For this reason, the exact dependence of the intensity operator in Eq. (4) on the frequency ω at $\mathbf{q}=0$ is used. However, we neglect in Eq. (4) the dependence of the quantity (5) on ω at $\mathbf{q}=0$, setting

$$\Delta G(\mathbf{p};0,\omega)\cong\mathrm{Im}\,G(p), \tag{15}$$

where

$$G(p)=G(p,\Omega+i0)=1/k_0^2-p^2-M(p),$$

$$M(p)=M(p,\Omega+i0).$$

Fraunhofer approximation

It is assumed that the imaginary part of the average Green's function in Eq. (15) is a "sharp" function of p corresponding to an approximation in the form of a Dirac δ-function

$$\operatorname{Im} G(p) \simeq -\frac{\pi}{2}\frac{\delta(p-k_{\text{eff}})}{k_0 n_{\text{eff}}}, \tag{16}$$

where the effective wave number k_{eff} is determined as the root of the equation $p^2-k_0^2+\operatorname{Re} M(p)=0$ and $n_{\text{eff}}=|k_{\text{eff}}/k_0||1+\partial \operatorname{Re} M(p)/\partial p^2|$ at $p=k_{\text{eff}}$. The approximation (16) is physically equivalent, according to Ref. 29, to the assumption[28] that, on the average, one effective scattering inhomogeneity of the medium is present in the Fraunhofer diffraction zone of another inhomogeneity, and it makes it possible to introduce the radiation intensity

$$I(\mathbf{s};\mathbf{q},\omega)=\frac{1}{(2\pi)^3}\int_0^\infty p^2 dp f(p\mathbf{s};\mathbf{q},\omega), \tag{17}$$

where $\mathbf{s}$ is a unit vector.

Low-density approximation

In the approximation of a low scatterer density n the mass operator and the intensity operator can be simply expressed (see, for example, Ref. 6) in terms of the scattering operator $T(\mathbf{r},\mathbf{r}';\Omega)=\tilde{T}(\Omega)\,\delta(\mathbf{r})\,\delta(\mathbf{r}')$ of an isolated scatterer, which is assumed to be a small, $k_0 r_0 \ll 1$, and is regarded as a resonant point scatterer of the model[27] with a scattering amplitude $\tilde{T}(\Omega)$ of the form

$$\tilde{T}(\Omega)=-\frac{4\pi k_0^2 r_{\text{eff}}}{k_r^2-k_0^2-ik_0^3 r_{\text{eff}}}. \tag{18}$$

Here the effective length $r_{\text{eff}}=1/\Lambda$ equals the reciprocal of the cutoff parameter $\Lambda\sim 1/r_0$ in the $\mathbf{k}$-space integration of the Fourier transform of the Green's function of the wave field in a homogeneous medium.[27] The resonance wave number k_r is determined together with the resonance frequency $\Omega_r=k_r C_0$ of a scatterer by the equation

$$k_r^2=\frac{r_{\text{eff}}}{r_0^3}\frac{3C_1^2}{C_0^2-C_1^2},$$

where C_1 is the speed of sound inside the scatterer. If the sound speed inside a scatterer is much less than the sound speed outside the scatterer, $C_1 \ll C_0$, then for $r_{\text{eff}}=r_0$ we find for the resonance frequency of the scatterer $\Omega_r\simeq\sqrt{3}C_1/r_0$. It can be verified by solving the problem of diffraction of a sound wave by a sphere (see, for example, Ref. 31) that for a small sphere ($k_0 r_0 \ll 1$) this value of Ω_r does indeed equal, to within a numerical factor, the lowest resonance frequency of the sphere, $(\pi/2)C_1/r_0$.

In direct proximity to a resonance of a point scatterer, $\Omega\simeq\Omega_r$ and $|\omega|\ll\Omega_r$, an approximation can be obtained for the two-frequency bilinear combination of scattering amplitudes (18) with a characteristic Lorentzian dependence on the frequency difference

$$\frac{\tilde{T}_1\tilde{T}_2}{|\tilde{T}|^2}\simeq\frac{1}{1-i\omega t_{\text{del}}}, \tag{19}$$

where the elementary delay time t_{del} mentioned in Sec. 1 assumes the value

$$t_{\text{del}}=\frac{2r_{\text{eff}}}{C_0}\left[4\left(\frac{\Omega-\Omega_r}{\Omega_r}\right)^2+\left(\frac{\Omega_r r_{\text{eff}}}{C_0}\right)^2\right]^{-1}. \tag{20}$$

Substituting the expressions (18) and (19) into the equations in Ref. 6 gives for the mass operator and the intensity operator the expressions $M(\mathbf{p},\Omega)=n\tilde{T}(\Omega)$ and

$$K(\mathbf{p},\mathbf{p}';\mathbf{q},\omega)=n\tilde{T}_1\tilde{T}_2\simeq\frac{4\pi/l}{1-i\omega t_{\text{del}}}, \tag{21}$$

where the mean free path l of the radiation is determined, as usual, by the equation $1/l=n|\tilde{T}|^2/4\pi$. From Eqs. (6) and (7), these equations for the mass operator and the identity operator lead in turn to the following representation of the quantity $A^+(\mathbf{p};\mathbf{q},\omega)$ in the low-density approximation, which takes account on the left-hand side of the kinetic equation (4) of the accumulation of acoustic energy inside the scatterers

$$A^+(\mathbf{p};\mathbf{q};\omega)\simeq-\frac{a}{1-i\omega t_{\text{del}}}. \tag{22}$$

The quantity a on the right-hand side of Eq. (22) equals the resonance efficiency factor in the formula $v_E=C_0^2/(1+a)C_{\text{ph}}$ for the radiation energy transport velocity, first introduced in Ref. 6. The quantity $C_{\text{ph}}=\Omega/k_{\text{eff}}$ denotes the phase velocity of a wave in a randomly-inhomogeneous medium. As emphasized in Ref. 6, the radiation energy transport velocity v_E in a randomly inhomogeneous resonant medium is different from both the phase velocity and the group velocity v_g, determined by the equation $C_0^2/C_{ph}v_g=1-n\partial\operatorname{Re}\tilde{T}/\partial E$ with $E=\Omega^2/C_0^2$. The Ward–Takahashi identity (12) in the low-density approximation gives at $\omega=0$ the following relation for the resonance efficiency factor a:

$$\frac{a}{n}=-\frac{\partial \operatorname{Re}\tilde{T}}{\partial E}+\frac{k_0}{2\pi}\operatorname{Im}\left(\frac{\partial\tilde{T}}{\partial E}\tilde{T}^*\right), \tag{23}$$

in agreement with the result of Ref. 27. The Lorentzian asymptotic expression (19) and the relation (21) with the mean free flight time t_l of the radiation determined in terms of the group velocity as $t_l=l/v_g$ make it possible to write the radiation energy transport velocity v_E as the Compton velocity[8,10] of excitation transfer $v_E=l/(t_{\text{del}}+t_l)$.

In the next section, in studying the solution of the modified radiation transport equation with delay, there appears another scattering time t_0 of radiation in a resonant medium, defined in terms of the phase velocity as $t_0=lC_{ph}/C_0^2$. The ratio of the elementary delay time (20) to the scattering time t_0 is related to the resonance efficiency factor a, as follows from Eq. (23), by the equation

$$\frac{t_{\text{del}}}{t_0}=a_{\text{kav}}=a+n\frac{\partial\operatorname{Re}\tilde{T}}{\partial E}, \tag{24}$$

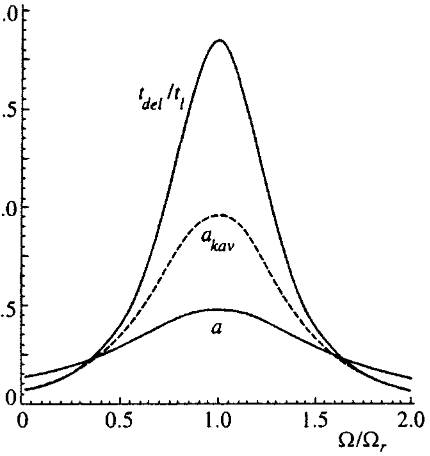


FIG. 1. Parameters a (bottom solid line), $a_{\rm kav}$ (dashed line), and ratio $t_{\rm del}/t_l$ (top solid line) versus the average frequency Ω of a pulse with $k_r r_0=1$, $\Lambda=1/r_0$, $C_0/C_1=2$, $\Omega_r=2C_1/r_0$, and $4\pi r_0^3 n/3=0.3$.

where $a_{\rm kav}$ was first introduced in Ref. 32. Comparing Eq. (24) with the definition of the group velocity v_g shows that $a_{\rm kav}<a+1$ for $v_g>0$.

Substitution of the expression (18) for the scattering amplitude $\widetilde{T}$ into the right-hand sides of Eqs. (23) and (24) and also into the equations for $1/l$, $C_{\rm ph}$, and v_g and simple calculations, similar to Ref. 5, give in direct proximity to resonance of a point scatterer the frequency dependences of the main physical parameters of the problem of pulse propagation in a randomly inhomogeneous medium consisting of resonance point scatterers, examples of which are given, for clarity, in Figs. 1 and 2 as a function of the ratio Ω/Ω_r of the frequency to its resonance value.

Modified radiative transfer equation with three Lorentzian delay kernels

We now turn to the kinetic equation (4). We multiply it by p^2 and integrate over p from $p=0$ to $p=\infty$, using the approximations (13)–(16), the definition (17), and also the expressions (21) and (22) for the intensity operator and the quantities that account for the accumulation of acoustic energy inside the scatterers. Together with the inverse Fourier transform in space $\mathbf{R}$ with respect to the variable $\mathbf{q}$ and the inverse Laplace transform in time t with respect to the variable ω, this leads to the following modified radiative transfer equation with three Lorentzian delay kernels:

$$\frac{C_{\rm ph}}{C_0^2}\frac{\partial}{\partial t}\left[I(\mathbf{s};\mathbf{R},t)+a\int_0^t\frac{dt'}{t_{\rm del}}\exp\left(-\frac{t-t'}{t_{\rm del}}\right)I(\mathbf{s};\mathbf{R},t')\right]$$
$$+(\mathbf{s}\nabla_{\mathbf{R}})I(\mathbf{s};\mathbf{R},t)=-\frac{1}{l}\int_0^t\frac{dt'}{t_{\rm del}}$$
$$\times\exp\left(-\frac{t-t'}{t_{\rm del}}\right)I(\mathbf{s};\mathbf{R},t')+\frac{1}{l}\int_0^t\frac{dt'}{t_{\rm del}}\int_{4\pi}\frac{d\mathbf{s}'}{4\pi}$$
$$\times\exp\left(-\frac{t-t'}{t_{\rm del}}\right)I(\mathbf{s}';\mathbf{R},t')+\frac{1}{(4\pi)^2}J(k_{\rm eff}\mathbf{s};\mathbf{R},t). \tag{25}$$

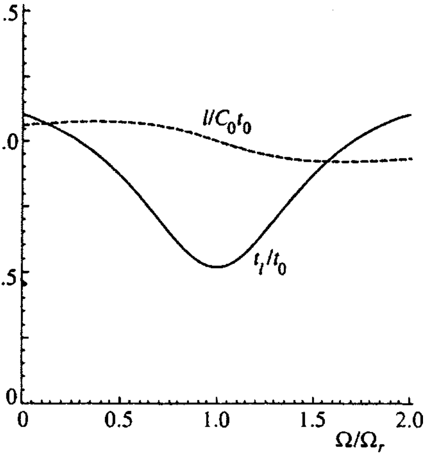


FIG. 2. Ratios l/C_0t_0 (broken curve) and t_l/t_0 (solid curve) versus the average pulse frequency Ω. The values of $k_r r_0$, Λ, C_0/C_1, Ω_r, and $4\pi r_0^3 n/3$ are the same as in Fig. 1.

Here $\mathbf{s}$ and $\mathbf{s}'$ are unit vectors and $d\mathbf{s}'$ is an element of solid angle. The expressions (8)–(10) for the ensemble-averaged energy quantities can be transformed, using the expressions (13)–(16), the definition (17), Eq. (22), and the above-mentioned inverse Fourier and Laplace transforms, to the following form in the space–time representation

$$\langle w(\mathbf{R},t)\rangle=\frac{1}{2\rho_0C_0^2}\int_{4\pi}d\mathbf{s}\int_0^t dt'\left[\delta(t-t')+\frac{a}{t_{\rm del}}\right.$$
$$\left.\times\exp\left(-\frac{t-t'}{t_{\rm del}}\right)\right]I(\mathbf{s};\mathbf{R},t'), \tag{26}$$

$$\langle S(\mathbf{R},t)\rangle=\frac{1}{2\rho_0C_0^2}\frac{C_0^2}{C_{\rm ph}}\int_{4\pi}d\mathbf{s}\;\mathbf{s}\;I(\mathbf{s};\mathbf{R},t), \tag{27}$$

$$\langle Q(\mathbf{R},t)\rangle=-\frac{1}{2\rho_0C_0^2}\frac{C_0^2}{C_{\rm ph}}\frac{1}{(4\pi)^2}\int_{4\pi}d\mathbf{s}\;J(k_{\rm eff}\mathbf{s};\mathbf{R},t). \tag{28}$$

Integrating the modified radiative transfer equation (25) over the variable $\mathbf{s}$ shows that its solution satisfies, on the basis (26)–(28), the ensemble-averaged Poynting theorem (11) in the temporal representation.

Modified radiative transfer equation with one Lorentzian delay kernel in the Sobolev approximation

A radiative transfer equation of the Sobolev equation type[15] with one Lorentzian delay kernel can be derived from the modified radiative transfer equation (25) by formally neglecting the delay effect in the expression in the square brackets on the left-hand side of Eq. (25), in the term with the accumulation of energy inside the scatterers, and the right-hand side of Eq. (25) in the first term with the extinction coefficient, but not in the second term with the scattering phase function. Moreover, the elementary delay time $t_{\rm del}$ must be neglected compared with the scattering time t_l on the right-hand side of the Compton representation of the radiation energy transport velocity v_E, thereby replacing it by the group velocity v_g.

4. SOLUTION OF THE ALBEDO PROBLEM OF DIFFUSE REFLECTION OF AN ACOUSTIC PULSE FROM A SEMI-INFINITE RESONANT RANDOMLY INHOMOGENEOUS MEDIUM

We shall now consider for the modified radiative transfer equation (25) with three Lorentzian delay kernels the problem of diffuse reflection of an acoustic pulse from a semi-infinite randomly inhomogeneous medium consisting of resonant point scatterers and occupying the half-space $z>0$. This problem presupposes finding a solution of the homogeneous modified radiative transfer equation (25) with the boundary condition

$$I(\mu;z=0,t)=\frac{1}{2}\,\delta(\mu-\mu_0)I_0(t),\tag{29}$$

where $\mu=s_z$ and $\mu_0=s_{0z}>0$ denote the z components of a unit vector $\mathbf{s}$ and unit vector $\mathbf{s}_0$ in the direction of propagation of the incident flattened pulse with a temporal profile given by the function $I_0(t)$. It is sufficient to find a particular solution of the albedo problem $I_\delta(\mu,\mu_0;z,t)$ with incidence of a pulse in the form of a Dirac δ-function with $I_0(t)=\delta(t)$, since the general solution is then obtained by forming the convolution of the solution found and $I_0(t)$.

Application of the principle of invariance for nonstationary problems

One could apply the principle of invariance, developed by Ueno[33] for the conventional nonstationary radiative transfer equation, in the time representation to solve the albedo problem for the modified radiative transfer equation (25) with the boundary condition (29). However, it is more convenient first to Laplace transform the modified radiative transfer equation (25) in time and then apply the Ambartsumyan principle of invariance[34] for the conventional stationary radiative transfer equation but now with the effective albedo of an elementary scattering event. This form of the principle of invariance for nonstationary problems was studied by Minin[19] in solving the modified radiative transfer equation with one Lorentzian delay kernel, as proposed by Sobolev.[15] After an inverse Laplace transform in time the Minin method leads to the following form of the solution of the albedo problem posed above for the modified radiative transfer equation (25) in the time representation:

$$I_\delta(-\mu,\mu_0;z=0,t)=\frac{1}{4\mu}\,\Sigma(\mu,\mu_0;t).\tag{30}$$

Here the scattering function $\Sigma(\mu,\mu_0;t)$ is expressed by means of the relation

$$t_0\Sigma(\mu,\mu_0;t)=\frac{\mu\mu_0}{\mu+\mu_0}\left[1+(1+a)\,\frac{\partial}{\partial\tilde{t}}+a_{\text{kav}}\,\frac{\partial^2}{\partial\tilde{t}^{\,2}}\right]\Sigma_1(\mu,\mu_0;\tilde{t}\,)\tag{31}$$

in terms of an intermediate scattering function $\Sigma_1(\mu,\mu_0;\tilde{t}\,)$, where the dimensionless time is $\tilde{t}=t/t_0$. The intermediate scattering function can be written as a convolution in time

$$\Sigma_1(\mu,\mu';\tilde{t}\,)=\int_0^{\tilde{t}}d\tilde{t}\,'H_1(\mu,\tilde{t}-\tilde{t}\,')H_1(\mu',\tilde{t}\,'),\tag{32}$$

by means of a generalized, time-dependent Chandrasekhar H-function $H_1(\mu,\tilde{t}\,)$ satisfying the equation

$$H_1(\mu,\tilde{t}\,)=\frac{e^{\nu_1\tilde{t}}-e^{\nu_2\tilde{t}}}{a_{\text{kav}}(\nu_1-\nu_2)}+\frac{1}{2}\,\mu\int_0^{\tilde{t}}d\tilde{t}\,'H_1(\mu,\tilde{t}-\tilde{t}\,')\times\int_0^1 d\mu'\,\frac{H_1(\mu',\tilde{t}\,')}{\mu+\mu'}.\tag{33}$$

The parameters ν_1 and ν_2 in the free term on the right-hand side of Eq. (33) are given by the equations

$$\nu_{1,2}=-\frac{1+a}{2a_{\text{kav}}}\left[1\mp\sqrt{1-\frac{4a_{\text{kav}}}{(1+a)^2}}\right],\tag{34}$$

where $\nu_1>\nu_2$ and it is assumed that $4a_{\text{kav}}<(1+a)^2$.

In the next two sections we shall examine the asymptotic values of the scattering functions $\Sigma(\mu;\mu_0;t)$ for small and large values of the dimensionless time $\tilde{t}$.

Case of small values $\tilde{t}\to 0$

In the limit of small values of the dimensionless time $\tilde{t}$ ($\tilde{t}\to 0$) the asymptotic solution of Eq. (33) can be approximated by the term on the right-hand side, setting $H_1(\mu,\tilde{t}\,)\to\tilde{t}/a_{\text{kav}}$. Substituting this value of $H_1(\mu,\tilde{t}\,)$ into the right-hand side of Eq. (32) gives $\Sigma_1(\mu,\mu';\tilde{t}\,)\to\tilde{t}^{\,3}/6a_{\text{kav}}^2$, which leads on the basis of Eq. (31) to an asymptotic expression of the scattering function for small values of the dimensionless time $\tilde{t}\to 0$

$$t_0\Sigma(\mu,\mu_0;t)\to\frac{\mu\mu_0}{\mu+\mu_0}\,\frac{\tilde{t}}{a_{\text{kav}}}.\tag{35}$$

This asymptotic expression is the single-scattering approximation for the leading edge of a diffusely reflected pulse.

Case of large values $\tilde{t}\to\infty$

In the limit of large values of the dimensionless time $\tilde{t}$ ($\tilde{t}\to\infty$) it is easier to use the asymptotic solution Eq. (33), Laplace transformed in time, in the corresponding limit of small values of the dimensionless frequency $\tilde{\omega}=\omega t_0$ ($\tilde{\omega}\to 0$), according to the idea, applied by Minin,[18] of finding for large values of the time the asymptotic solution of the modified radiative transfer equation with one Lorentzian delay kernel, proposed by Sobolev.[15] Minin's idea consists of using an asymptotic representation of the Chandrasekhar H-function[35] $H(\mu,\lambda)$ in the limit $\lambda\to 1$ of the form $H(\mu,\lambda)\to H(\mu,1)[1-\mu\sqrt{3(1-\lambda)}]$. In this method, after inverting the Laplace transform in time, the desired asymptotic representation of the scattering function for large values of the dimensionless time $\tilde{t}\to\infty$ is obtained. This asymptotic expression can be written as

$$t_0\Sigma(\mu,\mu_0;t)\rightarrow\frac{1}{2}\sqrt{\frac{3}{\pi}}H(\mu,1)H(\mu_0,1)\mu\mu_0\frac{\sqrt{1+a}}{\tilde{t}^{\,3/2}}\tag{36}$$

and represents the diffusion approximation for the trailing edge of a diffusely reflected pulse.

5. CONCLUSIONS

In this paper it was shown that a quasimonochromatic acoustic pulse can be trapped in a discrete randomly inhomogeneous medium consisting of an ensemble of randomly distributed scatterers when the carrying frequency of the pulse is close to the Mie resonance frequency of an isolated scatterer, and the elementary delay time can be greater than the mean free flight time of the radiation. Trapping of optical nonstationary radiation in gases whose atoms possess a resonance absorption line was first studied by Compton and Milne back in 1923 and 1926. In the last few years interest in the trapping of nonstationary radiation has actually reappeared in connection with attempts to interpret experiments on the observation of localization of classical, for example, acoustic, wave fields in discrete randomly-inhomogeneous media with a resonant elementary Mie scattering event. Specifically, as follows from the present work, the delay of the transfer of the radiation of classical wave fields in such media, which was discovered by an Amsterdam group[6] in 1991, is only a partial manifestation of the trapping of nonstationary radiation.

In this paper a systematic theory of trapping of nonstationary acoustic resonance radiation in a discrete randomly inhomogeneous medium was constructed on the basis of a new nonstationary radiative transfer equation with three Lorentzian delay kernels that takes account of the accumulation of acoustic energy inside scatterers and is consistent with the ensemble-averaged Poynting theorem for nonstationary acoustic radiation.

The albedo problem of diffuse reflection of a quasimonochromatic pulse from a semi-infinite resonant medium was solved for the nonstationary radiative transfer equation with three Lorentzian delay kernels which was derived. Simple asymptotic expressions were found in the single-scattering approximation and the diffusion approximation, respectively, for the scattering function for the leading and trailing edges of a diffusely reflected pulse, when the time scale of the incident pulse is short compared with the mean free flight time of the radiation and the incident pulse can be regarded as a Dirac δ-function. It was found that these asymptotic representations can be clearly stated for the scattering function of a pulse on the scale of the mean free flight time t_0, determined in terms of the phase velocity of the waves in a randomly inhomogeneous medium. On this scale the rise of the leading edge of the diffusely reflected pulse slows down as the parameter $a_{\mathrm{kav}}=t_{\mathrm{del}}/t_0$, equal to the ratio of the elementary delay time t_{del} to the mean scattering time t_0, increases. The dropoff of the trailing edge of a diffusely reflected pulse slows down as the parameter a characterizing the accumulation of acoustic energy inside scatterers increases.

This work was supported by the Russian Fund for Fundamental Research Grants Nos. 96-02-18829 and 97-02-17328.

[1] C. A. Condat and T. R. Kirkpatrick, in *Scattering and Localization of Classical Waves in Random Media*, edited by Ping Sheng, World Scientific, Singapore, 1990, Vol. 8.
[2] I. S. Graham, L. Piché, and M. Grant, Phys. Rev. Lett. **64**, 3135 (1990).
[3] S. John, Comments Condens. Matter Phys. **14**, 193 (1988).
[4] E. P. Wigner, Phys. Rev. **98**, 145 (1955).
[5] V. A. van Tiggelen, A. Lagendijk, M. P. van Albada, and A. Tip, Phys. Rev. B **45**, 12 233 (1992).
[6] M. P. van Albada, V. A. van Tiggelen, A. Lagendijk, and A. Tip, Phys. Rev. Lett. **66**, 3132 (1991).
[7] Yu. N. Barabanenkov and M. Yu. Barabanenkov, Phys. Lett. A **221**, 421 (1996).
[8] K. T. Compton, Philos. Mag. **45**, 750 (1923).
[9] E. A. Milne, J. Lond. Math. Soc. **1**, 40 (1926).
[10] T. Holstein, Phys. Rev. **72**, 1212 (1947).
[11] L. M. Biberman, Zh. Éksp. Teor. Fiz. **17**, 416 (1947).
[12] M. I. D'yakonov and V. I. Perel', Zh. Éksp. Teor. Fiz. **47**, 1483 (1964) [Sov. Phys. JETP **20**, 997 (1965)].
[13] V. M. Ermachenko, Zh. Éksp. Teor. Fiz. **51**, 1833 (1966) [Sov. Phys. JETP **24**, 1236 (1967)].
[14] Yu. A. Vdovin and V. M. Ermachenko, Zh. Éksp. Teor. Fiz. **54**, 148 (1968) [Sov. Phys. JETP **27**, 81 (1968)].
[15] V. V. Sobolev, *A Treatise on Radiative Transfer*, Van Nostrand, Princeton, N. J., 1963 [Russian original, Gostekhizdat, Moscow, 1956].
[16] I. N. Minin, in *Theory of Stellar Spectra* [in Russian], Nauka, Moscow, 1966, p. 159.
[17] S. A. Kaplan, S. F. Morozov, and L. V. Piskunov, Astrofizika **4**, 485 (1968).
[18] I. N. Minin, in *Theoretical and Applied Problems of Light Scattering* [in Russian], Nauka i tekhnika, Minsk, 1971, p. 59.
[19] I. N. Minin, Dokl. Akad. Nauk SSSR **154**, 1059 (1964) [Sov. Phys. Dokl. **9**, 114 (1964)].
[20] D. I. Nagirner, Astrofizika **5**, 31 (1969).
[21] V. V. Ivanov, *Transfer of Radiation in Spectral Lines* [in Russian], Nauka, Moscow, 1969 (National Bur. Standards, U.S. Govt. Printing Office, 1973).
[22] Yu. N. Barabanenkov and V. D. Ozrin, Phys. Lett. A **206**, 116 (1995).
[23] Yu. N. Barabanenkov, L. M. Zurk, and M. Yu. Barabanenkov, J. Electromagn. Waves Appl. **9**, 1393 (1995).
[24] Yu. N. Barabanenkov and V. D. Ozrin, Phys. Lett. A **154**, 38 (1991).
[25] Yu. N. Barabanenkov and V. D. Ozrin, Phys. Rev. Lett. **69**, 1364 (1992).
[26] J. Kroha, C. M. Soukoulis, and P. Wolfle, Phys. Rev. B **47**, 11093 (1993).
[27] T. M. Nieuwenhuizen, A. Lagendijk, and V. A. van Tiggelen, Phys. Lett. A **169**, 191 (1992).
[28] Yu. N. Barabanenkov and V. M. Finkel'berg, Zh. Éksp. Teor. Fiz. **53**, 978 (1967) [Sov. Phys. JETP **26**, 587 (1968)].
[29] Yu. N. Barabanenkov, Zh. Éksp. Teor. Fiz. **56**, 1262 (1969) [Sov. Phys. JETP **29**, 679 (1969)].
[30] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, Pergamon Press, N. Y. [Russian original, Nauka, Moscow, 1985].
[31] M. A. Isakovich, *General Acoustics* [in Russian], Nauka, Moscow, 1973.
[32] E. Kogan and M. Kaveh, Phys. Rev. B **46**, 10636 (1992).
[33] S. Ueno, J. Math. Anal. Appl. **4**, 1 (1962).
[34] V. A. Ambartsumyan, Dokl. Akad. Nauk SSSR **38**, 257 (1943).
[35] S. Chandrasekhar, *Radiative Transfer*, Clarendon Press, Oxford, 1950 [Russian translation, Inostr. Lit., Moscow, 1953].

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Interference phenomena in Doppler broadened quantum transitions: amplification of intense radiation without population inversion

A. K. Popov,*) V. M. Kuchin, and S. A. Myslivets

Institute of Physics, Siberian Branch of the Russian Academy of Sciences, Krasnoyarsk State University, Krasnoyarsk State Technical University, 660036 Krasnoyarsk, Russia

(Submitted 24 February 1997)
Zh. Éksp. Teor. Fiz. **113**, 445–470 (February 1998)

We study the effect of nonlinear interference processes on quantum transitions in intense resonant electromagnetic fields with allowance for changes in level populations, relaxation processes, incoherent excitation, and Doppler broadening of the transitions, on the absorption, amplification, and refraction of the interacting fields. The theory is generalized to the case of nonlinear interference interaction of two intense fields in open and closed three-level quantum systems. Using the density matrix, we derive general expressions in the case of stationary interaction that make it possible to analyze the optical characteristics for various configurations of the interfering transitions by a simple substitution of parameters. The possibility of amplifying light without saturated population inversion in a resonant transition is discussed. We formulate the conditions for such amplification and use examples to show that under appropriate changes in the initial level populations and the intensity of the auxiliary light, the inversionless amplification coefficient does not decrease with increasing intensity of the amplified radiation. We also show that allowance for these accompanying processes greatly affects the choice of optimal conditions for interference in optical transitions. As an illustration we list the results of a numerical analysis of possible experiments. © *1998 American Institute of Physics.* [S1063-7761(98)00502-2]

1. INTRODUCTION

Many ideas of quantum optics were formulated on the basis of Einstein's relations, which state that the probabilities of induced transitions involving photon absorption and emission are equal, and that they are related to the probability of spontaneous emission. From the classical viewpoint, emission and absorption of light are due to forced vibrations of coupled charges and the relationship between the phase of the radiation and the induced vibrations.

However, the same radiation can lead to several coherent interfering vibrations of varying nature. Interference can be either constructive or destructive, complete or incomplete, depending on phase and amplitude relationships. Here the corresponding optical response components can be amplified or suppressed. The macroscopic response of the medium is determined by quantum transitions in which the photons can participate in several processes simultaneously. In the presence of several resonant electromagnetic fields, the probability amplitudes of the quantum states contain several oscillating components at nearby frequencies. When we calculate the transition probabilities, in addition to squares of the absolute values of the corresponding components there are mixed terms that reflect the interference of the quantum transitions. Coherent nonlinear phenomena determined by the coevolution of quantum states mixed by an external field have become known as nonlinear interference effects.

In optics, wave interference leads to a redistribution of radiation in space (including zones with zero intensity) while the total radiative power is conserved. In quantum optics, nonlinear interference effects lead to a considerable change in the spectral characteristics of the transitions, so that unusual effects emerge. For instance, refractive indices whose signs change with frequency may emerge, which signals the onset in certain frequency ranges of amplification without population inversion on the levels involved in the transition, or, in contrast, the onset of absorption of radiation as a result of a transition with population inversion. The region where the sign changes corresponds to transparency, i.e., a lack of interaction between the amplified radiation and the resonant transition. The use of nonlinear interference effects makes it possible to capture and confine atoms in certain quantum states, a phenomenon known as coherent population trapping (CPT). On the basis of resonant nonlinear optical processes, entirely new methods of selective action on matter have been developed.

Here there is a profound analogy with classical ideas. Leaving aside the classification of elementary processes, which has been introduced for, and is true of, only the case of weak fields, we can predict and explain a broad spectrum of processes for which nonlinear interference effects are responsible. The theoretical bases of such an approach were developed by several groups of scientists in the USSR in the 1960s and 70s (see, e.g., Refs. 1–4). As noted earlier, one manifestation of such effects involving quantum transitions is amplification in a resonant transition without population inversion. The possibility of obtaining sign-varying contours of spectral lines in three-level systems in the microwave frequency range was demonstrated in theoretical papers (see, e.g., Ref. 5) and was corroborated experimentally (see, e.g., Ref. 6).

In several parameters the characteristics of optical transitions differ considerably from those of microwave transitions. The conditions for observing this effect in optical transitions for two-level systems was studied in Ref. 7, and for three-level systems in Refs. 2, 4 and 8, with the experimental aspects studied in Refs. 9–12. In Refs. 4 and 8, the criteria for amplification without population inversion at the center of the spectral line were derived, the evolution of the shape of the spectral line as the characteristics of incoherent and coherent pumping were varied was analyzed, and numerical estimates were based on an open V-scheme of transitions in neon. The anisotropy of the spectral lines and the distinctive features of nonlinear interference processes in Doppler-broadened transitions, with the strong-field intensity reaching values corresponding to Rabi frequencies of the order of the product of homogeneous and Doppler widths, were studied in Refs. 4 and 13. Lately there has been an upsurge of interest in the study of inversionless amplification in connection with recent advances in laser physics and new problems. The corresponding papers are cited in related collections of articles and in the latest works (see Ref. 14).

Nonlinear interference effects manifest themselves differently in different experimental situations. The phase relations and the amplitude of the interfering intratomic vibrations depend on the properties and relaxation characteristics of the transitions, the intensity of the fields, and the yields from resonances. For gaseous media, a characteristic feature of the typical experimental situations is the inhomogeneous Doppler broadening of the transitions. Depending on the value and sign of the Doppler shifts, the contributions to the macroscopic nonlinear polarization of the medium from the atoms moving with different velocities can amplify or suppress each other. The above effects are interrelated and can manifest themselves differently in absorption, refraction, and nonlinear optical mixing of radiation frequencies in macroscopic volumes. Their study is of considerable interest in both pure science and applications.

The present paper is a development of the results of Refs. 4, 8 and 13. We use a unified approach for the stationary case to study all possible open and closed three-level configurations (the V- and Λ-configurations) with allowance for different channels of relaxation and incoherent excitation to all levels, the inhomogeneous broadening of transitions, and the effect of the intensity of the amplified radiation on nonlinear interference and amplification without population inversion. We derive symmetric working equations that encompass these cases, and give numerical illustrations for characteristic examples of open and closed optical systems.

2. GENERAL EXPRESSIONS FOR THE REFRACTIVE INDEX, AMPLIFICATION, AND ABSORPTION COEFFICIENTS

We consider the interaction of two strong laser fields in a three level system. Figure 1 depicts the possible configurations of such systems: V, Λ, and ladder (H) configurations. We investigate the spectral characteristics of the refractive index and the amplification and absorption coefficients for the interacting fields with allowance for possible additional incoherent excitation of the levels (the excited atoms have a Maxwellian velocity distribution), as well as all kinds of processes of coherence and population relaxation in the system. Here one must distinguish between open and closed systems. In open systems, where the lower level is not the ground level, the rate of incoherent excitation of the levels by an external source is essentially independent of the induced transitions between the levels. On the other hand, in closed systems, where the lower level is the ground level, the rate of excitation of atoms with different velocities depends on the magnitude and velocity distribution of the population of the ground state, from which excitation to other levels starts, and in this way it depends on transitions in the system induced by electromagnetic fields.

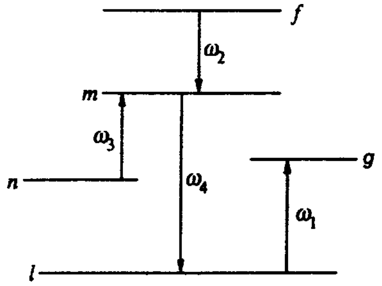


FIG. 1. Diagram of energy levels and transitions.

Below we study the spectral characteristics of a medium irradiated by light E_4 with tunable frequency ω_4 near the $l-m$ transition frequency. Depending on the configuration of the transitions being investigated, we turn on one of the strong fields, E_1, E_3, or E_2 with frequencies ω_1, ω_3, or ω_2 that are in resonance with the adjacent transitions depicted in Fig. 1. All fields are assumed to have the same polarization and are traveling waves that propagate in the positive or negative direction along the z axis:

$$E_j(z,t)=E_j \exp\{-i(\omega_j t-k_j z)\}+\text{c.c.},$$

where k_j can take negative values, and $j=1,2,3,4$.

2.1. Equations of the density matrix

We solve the problem by using a system of equations for the components of the density matrix ρ of a mixed quantum mechanical ensemble. These equations make it possible to allow (in a phenomenological manner) for processes of incoherent excitation and relaxation that are due to the interaction with other systems whose properties do not depend on the ensemble in question. In operator form and in the interaction picture, these equations are

$$\frac{d\rho}{dt}=-\frac{i}{\hbar}[V,\rho]+R+q, \tag{1}$$

where V is the Hamiltonian describing the interaction of the atom and external electromagnetic radiation, and R and q describe relaxation processes and incoherent excitation of the levels.

Only stationary states will be considered. Here the form of Eq. (1) suggests that since the amplitude of the interaction

Hamiltonian is time-independent, solutions for ρ_{ij} for the interaction with a plane wave will take the form

$$\rho_{ii}=r_i, \quad \rho_{ij}=r_{ij}\exp[i(\Omega t-kz)], \quad \rho_{ji}=\rho_{ij}^*,$$

where the labels correspond to a transition from a lower level to a higher level.

The equation for the diagonal elements of the density matrix ρ have the form of balance equations for the mean level populations. Hence, if there is relaxation in the system and incoherent pumping takes place, only the diagonal elements of q are nonvanishing, and the expressions for these elements can be written, on empirical grounds, as the product of the rate (probability per unit time) of incoherent excitation to a given level and the population of the initial level. For each configuration of the three-level system below, we consider two cases, an open configuration and a closed configuration, whose features were described earlier. We call the level populations in the absence of radiation the initial level populations. Each relaxation term can be written as a product of the appropriate relaxation constants and an element of the density matrix. In addition to the fact that the equations for the density matrix make it possible to allow for these statistical processes, in the stationary case they make it possible to replace the system of differential equations by a system of algebraic equations.

Below we employ notation in which double indices for off-diagonal elements, which indicate transitions from a lower level to a higher one, are replaced by the transition indices used in Fig. 1. For instance,

$$L_{nn}\rho_{nn}=q_n-i[V,\rho]_{nn}+\gamma_{mn}\rho_{mm}, \quad L_{nn}=\frac{d}{dt}+\Gamma_n,$$

$$L_{lm}\rho_{lm}=L_4\rho_4=-i[V,\rho]_{lm}, \quad L_{lm}=\frac{d}{dt}+\Gamma_{lm},$$

$$V_{lm}=G_{lm}\cdot\exp\{i[\Omega_4 t-kz]\}, \quad G_{lm}=-\frac{\mathbf{E}\cdot\mathbf{d}_{lm}}{2\hbar}, \tag{2}$$

where $\mathbf{d}_{lm}$ is the dipole moment of the transition, $q_n=\Sigma_j w_{nj}r_j$, with w_{nj} the rate (probability per unit time) of excitation of atoms into state n from lower levels; γ_{mn} are the rates of relaxation transitions between the levels; Γ_{mn} are the homogeneous transition halfwidths; $\Gamma_n=\Sigma_j\gamma_{\gamma_{nj}}$ are the reciprocal lifetimes of the levels; and $\Omega_4=\omega_4-\omega_{ml}$ is the offset from resonance. In the absence of collisions, $\Gamma_{ij}=(\Gamma_i+\Gamma_j)/2$.

In this notation the equations for the absorption (amplification) coefficient $\alpha_4(\Omega_4)$ and the refractive index $n_4(\Omega_4)$ assume the form (if we use the solutions for the off-diagonal elements of the density matrix)

$$\frac{\alpha_4(\Omega_4)}{\alpha_4^0}=\mathrm{Re}\,\frac{r_4/G_4}{(r_4/G_4)_0}, \tag{3}$$

$$\frac{n_4(\Omega_4)-n_4^{\mathrm{nr}}}{2(n_{4\,\mathrm{max}}^0-n_4^{\mathrm{nr}})}=\mathrm{Im}\,\frac{r_4/G_4}{(r_4/G_4)_0}. \tag{4}$$

Here α_4^0 is the value of the coefficient in vanishingly weak fields at the center of the line, n_4^{nr} represents the contribution of all nonresonant levels to the refractive index at the given frequency (in gases at low pressures, $n_4^{\mathrm{nr}}\approx 1$), $n_{4\,\mathrm{max}}^0$ is the maximum value of the refractive index in the resonance region in the zero-field limit, and $(r_4/G_4)_0$ is the ratio at resonance in zero fields.

If the atom's velocity is v, we must add $\mathbf{v}\cdot\boldsymbol{\nabla}\rho$ to the left-hand side of Eq. (1). In the cases studied in this paper this proves to be equivalent to allowing for a Doppler shift in the resonances, i.e., to replacing the offset from resonance Ω_j by $\Omega_j'=\Omega_j-k_j v$ in the final expression. In what follows, the prime on the offset is dropped, so that we assume that the Doppler shift has been taken into account. We also use the notation $\Delta r_4=r_l-r_m$ and $\Delta n_4=n_l-n_m$, with n_i the level populations in the absence of fields; these are described by $n_i=q_i/\Gamma_i+(\gamma_{ki}/\Gamma_i)(q_k/\Gamma_k)$.

In open configurations, the population of the ground state is usually higher than the other populations. Hence, incoherent excitation from this level plays the leading role, and q_i can be assumed to be constant even in the presence of fields.

The distinctive features of closed configurations can be taken into account by replacing one of the equations for level populations (say, the equation for the lower level) by the law of conservation of the sum of populations. We also use the notation $P_i=\Gamma_i+i\Omega_i$ for the resonance denominators (e.g., $P_{lm}=P_4=\Gamma_4+i\Omega_4$, $P_{lm}=P_{ml}^*$, $P_{lf}=P_{42}=\Gamma_{lf}+i(\Omega_4+\Omega_2)$, etc.).

We next study the resonant nonlinear interaction of two strong light fields in three-level quantum systems with different configurations of the correlated transitions.

2.2. Generalized equations describing nonlinear interference effects in the field of two intense radiative fields for various transition configurations

The calculations in the Appendices show that although there are important difference in the manifestations of the interference effects in quantum transitions of various configurations (open and closed), the equations for the refractive index, the amplification and absorption coefficients, and the populations of the energy levels—the latter being established by various coherent and incoherent processes—can be presented in a systematic manner for all cases depicted in Fig. 1:

$$\frac{\alpha_4}{\alpha_{04}}=F_1\frac{\Delta r_4}{\Delta n_4}\mp F_2\frac{\Delta r_i}{\Delta n_4}, \quad \frac{\alpha_i}{\alpha_{0i}}=F_3\frac{\Delta r_i}{\Delta n_i}\mp F_4\frac{\Delta r_4}{\Delta n_i}, \tag{5}$$

$$\frac{n_4(\Omega_4)-n_4^{\mathrm{nr}}}{2(n_{4\,\mathrm{max}}^0-n_4^{\mathrm{nr}})}=f_1\frac{\Delta r_4}{\Delta n_4}\mp f_2\frac{\Delta r_i}{\Delta n_4},$$

$$-\frac{n_i(\Omega_i)-n_i^{\mathrm{nr}}}{2(n_{i\,\mathrm{max}}^0-n_i^{\mathrm{nr}})}=f_3\frac{\Delta r_i}{\Delta n_i}\mp f_4\frac{\Delta r_4}{\Delta n_i}. \tag{6}$$

Here and in what follows, the label i indicates a transition in resonance with the auxiliary applied field (see Fig. 1), and

$$F_1+if_1=\frac{\Gamma_4}{P_4}\frac{1+u_2}{1+g_1+u_2}, \quad F_2+if_2=\frac{\Gamma_4}{P_4}\frac{g_2}{1+g_1+u_2},$$

$$F_3+if_3=\frac{\Gamma_i}{P_i^*}\frac{1+g_1}{1+g_1+u_2}, \quad F_4+if_4=\frac{\Gamma_i}{P_i^*}\frac{u_1}{1+g_1+u_2}, \tag{7}$$

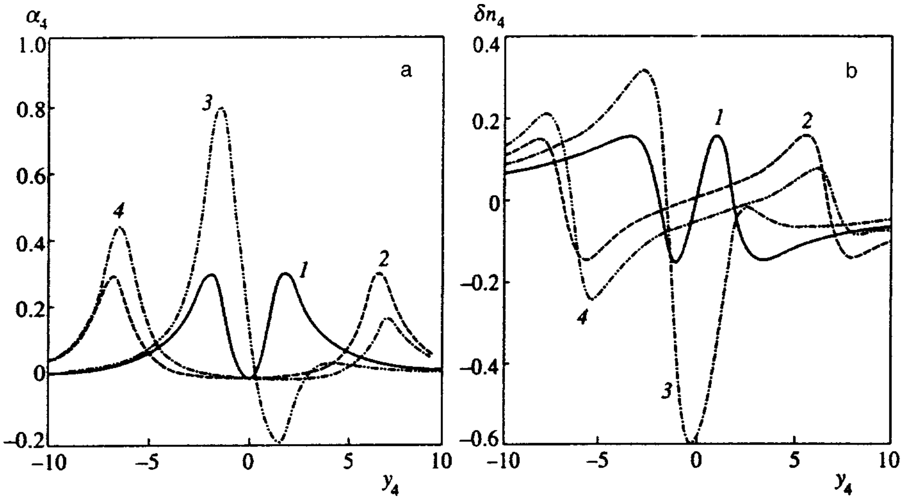


FIG. 2. Dependence of the absorption coefficient (a) and refractive index (b) in the homogeneously broadened transition $l-m$ (l is the excited state) in the presence of a strong field in the adjacent transition $l-g$ on the normalized offset $y_4=\Omega_4/\Gamma_{lm}$ ($S_4=0$): curves *1*, $S_1=5$ and $\Omega_1=0$; curves *2*, $S_1=100$ and $\Omega_1=0$; curves *3*, $S_1=5$ and $\Omega_1=2\Gamma_{lg}$; curves *4*, $S_1=100$ and $\Omega_1=2\Gamma_{lg}$.

$$g_1=\frac{|G_i|^2}{P_4P_{4i}},\quad g_2=\frac{|G_i|^2}{P_i^*P_{4i}},\quad u_1=\frac{|G_4|^2}{P_4P_{4i}},\quad u_2=\frac{|G_4|^2}{P_i^*P_{4i}}. \tag{8}$$

The functions F_j have the same form for all schemes except the H-scheme. In the H-scheme, we must remove the conjugation sign on P_i in the denominators of the expressions for g_2, u_2, F_3, and F_4.

The population differences, which depend on the radiative intensities, can also be written in general form:

$$\Delta r_4=\frac{\Delta n_4X_2\mp\Delta n_iX_3}{X_1X_2-X_3X_4},\quad \Delta r_i=\frac{\Delta n_iX_1\mp\Delta n_4X_4}{X_1X_2-X_3X_4}. \tag{9}$$

The minus sign in Eqs. (5), (6), and (9) refers to the V(E_4,E_1)- and Λ(E_4,E_3)-schemes, and the plus sign, to the H(E_4,E_2)-scheme. Furthermore,

$$X_1=1+a_{14}\kappa_4F_1+a_{1i}\kappa_iF_4,\quad X_2=1+a_{24}\kappa_4F_2+a_{2i}\kappa_iF_3,$$

$$X_3=a_{34}\kappa_4F_2+a_{3i}\kappa_iF_3,\quad X_4=a_{44}\kappa_4F_1+a_{4i}\kappa_iF_4,$$

$$a_{14}=-a_{34},\quad a_{1i}=-a_{3i},\quad a_{24}=-a_{44},\quad a_{2i}=-a_{4i}, \tag{10}$$

with κ_4 and κ_i the saturation parameters for the transitions 4 and i, respectively. For open configurations,

$$\kappa_4=\frac{2|G_4|^2(\Gamma_l+\Gamma_m-\gamma_4)}{\Gamma_l\Gamma_m\Gamma_4}, \tag{11}$$

and the saturation parameters κ_i as well as the parameters a_{mn}, which depend only on the relaxation constants, are given below for each scheme.

2.2.1. *V-scheme (fields E_4 and E_1; $i=1$)*

Open configuration:

$$\kappa_i=\kappa_1=\frac{2|G_1|^2(\Gamma_g+\Gamma_l-\gamma_1)}{\Gamma_g\Gamma_l\Gamma_1},\quad a_{2i}=a_{21}=1,\quad a_{14}=1,$$

$$a_{3i}=a_{31}=\frac{\Gamma_g-\gamma_1}{\Gamma_g+\Gamma_l-\gamma_1},\quad a_{44}=\frac{\Gamma_m-\gamma_4}{\Gamma_l+\Gamma_m-\gamma_4}. \tag{12}$$

Closed configuration:

$$\kappa_4=\frac{4|G_4|^2}{\Gamma_m\Gamma_4},\quad \kappa_i=\kappa_1=\frac{4|G_1|^2}{\Gamma_g\Gamma_1},$$

$$a_{3i}=a_{31}=0.5\Delta n_4,\quad a_4=0.5\Delta n_1,$$

$$a_{2i}=a_{21}=0.5(1+\Delta n_1),\quad a_{14}=0.5(1+\Delta n_4). \tag{13}$$

2.2.2. *Λ-scheme (fields E_4 and E_3; $i=3$)*

Open configuration:

$$\kappa_3=\frac{2|G_3|^2(\Gamma_m+\Gamma_n-\gamma_3)}{\Gamma_m\Gamma_n\Gamma_3},\quad a_{2i}=a_{23}=1,\quad a_{14}=1,$$

$$a_{3i}=a_{33}=\frac{\Gamma_n}{\Gamma_l}\,\frac{\Gamma_l-\gamma_4}{\Gamma_m+\Gamma_n-\gamma_3},\quad a_{44}=\frac{\Gamma_l}{\Gamma_n}\,\frac{\Gamma_n-\gamma_3}{\Gamma_m+\Gamma_l-\gamma_4}. \tag{14}$$

Closed configuration:

$$\kappa_4=\frac{4|G_4|^2}{\Gamma_m\Gamma_4},$$

$$a_{3i}=a_{33}=1+\Delta n_4-(1+2\Delta n_4)\,\frac{\Gamma_m-\gamma_3}{\Gamma_m+\Gamma_n-\gamma_3},$$

$$a_{44}=0.5\left[1-\frac{\gamma_3}{\Gamma_n}+\Delta n_3\left(1+\frac{\gamma_3}{\Gamma_n}\right)\right],$$

$$a_{2i}=a_{23}=1+\Delta n_3\,\frac{\Gamma_n-(\Gamma_m-\gamma_3)}{\Gamma_n+\Gamma_m-\gamma_3},$$

$$a_{14}=0.5\left[1+\Delta n_4\left(1+\frac{\gamma_3}{\Gamma_n}\right)\right]. \tag{15}$$

2.2.3. *H-scheme (fields E_4 and E_2; $i=2$)*

Open configuration. As noted earlier, in the functions F_3, F_4, g_2, and u_2, we must remove the conjugation sign on P_i:

$$\kappa_2=\frac{2|G_2|^2(\Gamma_f+\Gamma_m-\gamma_2)}{\Gamma_f\Gamma_m\Gamma_2},\quad a_{14}=1,\quad a_{2i}=a_{22}=1,$$

$$a_{3i}=a_{32}=\frac{\Gamma_l-\gamma_4}{\Gamma_l}\,\frac{\Gamma_f-\gamma_2}{\Gamma_m+\Gamma_f-\gamma_2},\quad a_{44}=\frac{\Gamma_l}{\Gamma_l+\Gamma_m-\gamma_4}. \tag{16}$$

Closed configuration:

$$\kappa_4=\frac{4|G_4|^2}{\Gamma_m\Gamma_4},$$

$$a_{3i}=a_{32}=(1+2\Delta n_4)\,\frac{\Gamma_f-\gamma_2}{\Gamma_m+\Gamma_f-\gamma_2}-\Delta n_4,$$

$$a_{44}=0.5(1-\Delta n_2),\quad a_{14}=0.5(1+\Delta n_4),$$

$$a_{2i}=a_{22}=1+\Delta n_2\,\frac{\Gamma_m-(\Gamma_f-\gamma_2)}{\Gamma_m+\Gamma_f-\gamma_2}. \tag{17}$$

3. GENERAL ANALYSIS OF THE EQUATIONS

The emitted power per unit volume due to induced emission, W_i, and the absorbed power W_a, are directly proportional to the number of atoms per unit volume that are capable, respectively, of emitting or absorbing radiation (the populations of the upper and lower levels), and proportional to the probability $b_i(I)$ of induced emission and the probability $b_a(I)$ of absorption generated by radiation of intensity I. According to the Einstein relations, these probabilities are usually assumed equal, which means that population inversion is needed for amplification of radiation.

To make amplification without population inversion possible, the probability of induced emission must exceed the probability of absorption. This is made possible by nonlinear interference effects. This introduces a situation in which the absorption of light by atoms in the lower energy levels decreases and the probability of induced emission of light by atoms in the higher levels increases or remains unchanged. As noted earlier, a difference between the probabilities of induced emission and absorption due to nonlinear interference effects was demonstrated in Refs. 2–4, 13, 15–17; the conditions for the onset of amplification of light without population inversion due to such effects were studied for two-level optical systems in Ref. 7 and for three-level systems, in Refs. 4 and 8.

Let us examine, for instance, the interaction of fields E_3 and E_4 in the Λ-scheme (see Fig. 1). Quantum interference can be achieved when the transitions involve a coherent superposition of states that are close in energy. Such states may be real sublevels (this situation was considered, for instance, in Ref. 1). A similar situation can be realized for quasilevels through the use of electromagnetic radiation. For example, in the case of exact resonance, the lower level l in the V-scheme can be represented by a coherent superposition of two symmetrically disposed quasilevels l_1 and l_2 produced by a resonant external field. The light-emitting atoms can transit from level m to either level l_2 or l_2, and the processes are independent. If the atom in the lower level is in a coherent superposition of states, transitions from these states are not independent and interfere. Under certain conditions, the interference is destructive and the absorption process is suppressed.

In some cases, interfering frequency-degenerate oscillations within a single atom, induced by auxiliary radiation, can be interpreted as being due to various correlated quantum transitions, all contributing to the same process. These might be, for example, one- and two-photon contributions to an optical process related to emission or absorption at the given frequency. For instance, in the V-scheme (Fig. 1), photons can be observed as a result of transitions from level l to level m and as a result of transitions from level g to level m with the simultaneous participation of two photons, one with energy $\hbar\omega_1$ emitted in such transitions, and the other, with energy $\hbar\omega_4$, being absorbed.

The classification of quantum transitions as one- and multiphoton transitions was introduced on the basis of the difference in their frequency-correlation properties in the context of perturbation theory. In strong resonant electromagnetic fields these properties undergo dramatic changes, so that the initial classification becomes meaningless from the standpoint of physics.[2–4,15,16] The corresponding variations manifest themselves in resonant emission and absorption spectra; for instance, the conditions for resonances free of Doppler broadening change dramatically.[2–4,16,18]

Here it is much more convenient to use a classification scheme based on the structure of the solution for the off-diagonal elements of the density matrix.[2–4,13] In this classification scheme, many experimental results can explained and predicted by using the concept of interfering components in off-diagonal density-matrix elements, components induced by the interacting fields. The amplitudes of these components can be varied by changing the intensities of the respective fields, while the phases can be varied by changing the magnitudes and signs of their offset from one- and multiphoton resonances.

For instance, let us turn to the case, considered in Refs. 4 and 8, of inversionless amplification of weak radiation in the presence of a strong field in resonance with the adjacent transition in the V-configuration. According to Eq. (A1) for ρ_4 (see Appendix A), polarization at this frequency consists of two interfering components, one of which is due to the coherence ρ_{41} jointly induced by the probe and strong light fields, i.e., correlated transitions $g-l$ and $l-m$ between the levels g and m. As a result, the solution (A2) for the polarization with $|u_{1,2}|\ll 1$,

$$\frac{r_4}{iG_4}=\frac{\Delta r_4-\Delta r_1(|G_1|^2/P_1^*P_{41})}{P_4[1+(|G_1|^2/P_4P_{41})]}, \tag{18}$$

makes it possible to identify three effects of the auxiliary strong field, and each can be observed individually or in conjunction with the other two:

1) The effect of moving populations (the field-dependence of Δr_4 and Δr_1) manifests itself in strong coherence relaxation ($\Gamma_{41}\gg\Gamma_{4,1}$) and in Doppler-broadened transitions for counterpropagating waves in the V- and Λ-schemes (for unidirectional waves in the H-scheme).

2) The denominator in (18) has two roots with respect to Ω_4 (one- and two-photon resonances modified by the field). This is a reflection of resonance splitting in this transition, which can show up separately at $n_g=n_l$.

3) The second term in the numerator, taken together with the denominator, reflects the presence of a nonlinear interference effect, since it influences only the shape of the spectral line and not the total intensity. The integral of the second term with respect to Ω_4 (along with the denominator) is zero, and the integral of the first term with respect to Ω_4 (along with the denominator) is equal to Δr_4.

Thus, the variation in the total absorption or amplification is determined by the variation in the population difference, while the nonlinear interference effects only influence the shape of the spectral line. In the expansion in $|G_1|^2/P_4P_{41}$ in weak fields, such effects show up independently, and under some conditions (e.g., for Doppler-broadened short-wave transitions, $\omega_4 > \omega_i$) they can be observed in pure form. These effects are associated with the possibility of changes in sign of the amplification coefficient for $\Delta r_1(|G_1|^2/\Gamma_1\Gamma_{41}) > \Delta r_4$ without changes in the sign of Δr_4. This means that both amplification without population inversion and electromagnetically induced transparency (the fact that absorption vanishes at finite values of Δr_4) can occur in this system. Here, under resonant conditions, the effect is not tantamount to amplification due to inversion population in two-photon transitions. A component proportional to the population difference in the two-photon transition is produced only if the yield of the one-photon resonance is high, whereupon the interference between ladder and two-photon quantum transitions vanishes. Note that this component is produced by field-dependent factors in both the numerator and denominator of Eq. (18). At $|\Omega_1| \approx |\Omega_4| \gg \Gamma_1, \Gamma_4$, $|g_{1,2} \ll 1|$, $P_4 \approx i\Omega_4$, and $P_1 \approx i\Omega_1 \approx i\Omega_4$, Eq. (18) assumes the form

$$\begin{aligned}\frac{\alpha(\Omega_4)}{\alpha^0(0)} &\approx \frac{\Gamma_4^2\Delta r_4}{\Omega_4^2\Delta n_4} - \mathrm{Re}\,\frac{\Gamma_4(\Delta r_4 g_1 + \Delta r_1 g_2)}{i\Omega_4\Delta n_4} \\ &\approx \frac{\Gamma_4^2\Delta r_4}{\Omega_4^2\Delta n_4} - \frac{\Gamma_4\Gamma_{14}}{\Gamma_{14}^2 + (\Omega_4 - \Omega_1)^2}\,\frac{|G_1|^2(\Delta r_1 - \Delta r_4)}{\Omega_4^2\Delta n_4} \\ &= \frac{\Gamma_{lm}^2(r_l - r_m)}{(n_l - n_m)\Omega_4^2} - \frac{\Gamma_{gm}\Gamma_{lm}}{\Gamma_{gm}^2 + (\Omega_4 - \Omega_1)^2}\,\frac{|G_1|^2(r_m - r_g)}{\Omega_4^2(n_l - n_m)}.\end{aligned} \tag{19}$$

The last two terms in this equation describe a two-photon process, and originate in the second term in the numerator (together with the denominator) in Eq. (18). Thus, the lack of interference leads to the need for population inversion in the initial and final states ($r_m = n_m > r_g$) in order to amplify the probing field.

Let us examine the conditions for inversionless amplification of intense radiation, which are related directly to the problem of an inversionless laser. For the V- and Λ-schemes, using Eqs. (5) for each transition we have, respectively,

$$\frac{\Delta r_4}{\Delta r_i} < \frac{F_2}{F_1}, \quad \frac{\Delta r_i}{\Delta r_4} < \frac{F_4}{F_3}. \tag{20}$$

These conditions simplify at the center of the line

$$\frac{\Delta r_4}{\Delta r_i} < \frac{g_2}{1 + u_2}, \quad \frac{\Delta r_i}{\Delta r_4} < \frac{u_1}{1 + g_1}. \tag{21}$$

In the H-scheme, the conditions for amplification without population inversion are

$$\frac{\Delta r_4}{\Delta r_i} < -\frac{F_2}{F_1}, \quad \frac{\Delta r_i}{\Delta r_4} < -\frac{F_4}{F_3}. \tag{22}$$

This implies that in contrast to the V- and Λ-schemes, there must either be population inversion of the levels involved in the adjacent transition, or amplification occurs under certain conditions in the wings of the resonance.

Level populations also depend on the intensities and frequencies of both fields and on the initial level populations. Thus, optimization should be carried out in a self-consistent manner. Below, the corresponding dependence is illustrated by numerical examples.

We now formulate the general criteria for optimal conditions for the center of the line. To be definite, we examine a transition scheme with a common lower level, i.e., an open V-configuration.

Using Eqs. (5) and (9), we obtain

$$\begin{aligned}\frac{\alpha_4}{\alpha_{04}} &= A\left\{1 - \frac{\Delta n_1}{\Delta n_4}\,\frac{|G|_1^2}{\Gamma_1^2(1 + \kappa_1 + |G|_4^2/\Gamma_1\Gamma_{41})}\right. \\ &\qquad \left.\times\left[\left(1 - \frac{\gamma_1}{\Gamma_g}\right)\frac{2\Gamma_1}{\Gamma_l} + \frac{\Gamma_1}{\Gamma_{41}}\right]\right\}, \\ A &= \frac{1 + \kappa_1 + |G|_4^2/\Gamma_1\Gamma_{41}}{(1 + g_1 + u_2)(X_1X_2 - X_3X_4)}.\end{aligned} \tag{23}$$

If the amplified fields are negligible, this expression transforms into the one studied in Refs. 4 and 8:

$$\begin{aligned}\frac{\alpha_4}{\alpha_{04}} &= \frac{1}{1 + |G|_1^2/\Gamma_4\Gamma_{41}}\left\{1 - \frac{\Delta n_1}{\Delta n_4}\,\frac{|G|_1^2}{\Gamma_1^2(1 + \kappa_1)}\right. \\ &\qquad \left.\times\left[\left(1 - \frac{\gamma_1}{\Gamma_g}\right)\frac{2\Gamma_1}{\Gamma_l} + \frac{\Gamma_1}{\Gamma_{41}}\right]\right\}.\end{aligned} \tag{24}$$

If we compare (23) and (24), we see that all the main effects of a strong field in resonance with the adjacent transition are retained, and that there are additional effects associated with the amplified radiation.

Equations (20)–(24) make it possible to interrelate the main parameters of the inversionless amplification region. In resonance, the population differences between the interacting transitions for the V-scheme can be described by Eqs. (9) (see also Appendix A):

$$\begin{aligned}\Delta r_4 &= r_l - r_m = \frac{\Delta n_4X_2 - \Delta n_1X_3}{X_1X_2 - X_3X_4}, \\ \Delta r_1 &= r_l - r_g = \frac{\Delta n_1X_1 - \Delta n_4X_4}{X_1X_2 - X_3X_4}, \\ X_1 &= 1 + \frac{\kappa_4[1 + u_2]}{1 + g_1 + u_2} - a_1\,\frac{\kappa_1u_1}{1 + g_1 + u_2}, \\ X_2 &= 1 + \frac{\kappa_1[1 + g_1]}{1 + g_1 + u_2} - a_4\,\frac{\kappa_4g_2}{1 + g_1 + u_2},\end{aligned} \tag{25}$$

$$X_3=a_1\,\frac{\kappa_1[1+g_1]}{1+g_1+u_2}-\frac{\kappa_4 g_2}{1+g_1+u_2},$$

$$X_4=a_4\,\frac{\kappa_4[1+u_2]}{1+g_1+u_2}-\frac{\kappa_1 u_1}{1+g_1+u_2}.$$

There is no inversion of saturated populations if

$$\frac{\Delta n_1}{\Delta n_4}<\frac{X_2}{X_3},\quad \frac{\Delta n_4}{\Delta n_1}<\frac{X_1}{X_4}.$$

Since the constants

$$a_i=a_1=\frac{\Gamma_g-\gamma_1}{\Gamma_g+\Gamma_l-\gamma_1},\quad a_4=\frac{\Gamma_m-\gamma_4}{\Gamma_l+\Gamma_m-\gamma_4}$$

are always less than unity, the ratios X_2/X_3 and X_1/X_4 are greater than unity for all values of the saturation parameters of both fields. This means that if the ratio of the initial level populations, $\Delta n_1/\Delta n_4$, does not exceed unity, there is not a single saturation parameter or relaxation constant at which inversion of saturated populations occurs.

Furthermore, there is no population inversion if the above constants obey the following inequalities (see (25)):

$$a_1<\frac{\kappa_4 g_2}{\kappa_1(1+g_1)},\quad a_4<\frac{\kappa_1 u_1}{\kappa_4(1+u_2)}.$$

Other cases require special treatment.

The threshold and laser power are dictated by

$$\alpha_4=T, \tag{26}$$

where T represents the radiative loss per pass of the cavity per unit length of the gain medium. Thus, the above expressions determine the conditions for, and characteristics of, inversionless lasing.

4. NUMERICAL ANALYSIS

We now illustrate the above behavior of open systems using neon atoms, and of closed systems, using a model with parameters close to the transitions of atomic sodium. For neon, we select the same transitions that were used in Refs. 4 and 8 to analyze the conditions for weak-field inversionless amplification involving homogeneously broadened transitions: the $2s_2-2p_4$ transition for the long-wave field (gl), and the $3s_2-2p_4$ transition for the short-wave field (ml). The relaxation constants for these transitions are well known: $\Gamma_m=3\times10^7$ s^{-1}, $\Gamma_l=5\times10^7$ s^{-1}, $\Gamma_g=10^7$ s^{-1}, and $\gamma_{ml}=\gamma_{gl}=0.5\times10^7$ s^{-1}. Equations for the velocity-averaged absorption coefficients of a weak field at the high-frequency transition in the presence of a strong field with the characteristic Rabi frequency not exceeding the Doppler width of the low-frequency transition were obtained in Refs. 4 and 13, where it was established that the line profile of the probe field is anisotropic, and the authors discussed the differences among the saturation effects, the conditions needed for separate manifestations of each such effect, and the modification that the effects undergo with increasing strong-field intensity. The same questions were studied for population motion and the nonlinear interference effect. Moreover, differences in the spectral manifestations of these effects were investigated, and the various features of lasing under coherent pumping were analyzed. Hence, we focus below on the effects associated with an increase in the intensity of the amplified radiation.

Figures 2–8 show the results of numerical analysis done by using the equations of Sec. 2.2 and the Appendices. In averaging over velocity, the initial level population was written

$$n_j=\frac{N_j}{\sqrt{\pi}\bar{v}}\exp\left[-\left(\frac{v_z}{\bar{v}}\right)^2\right],$$

where N_j are the velocity-integrated populations. In Figs. 2–8 we use the notation $\alpha_i=\alpha_i(\Omega_i)/\alpha_i^0$ for the normalized absorption coefficient (positive values) and amplification coefficient (negative values), and $\delta n_i=[n_i(\Omega_i)-n_i^{\mathrm{nr}}]\times[2(n_{i\,\max}^0-n_i^{\mathrm{nr}})]^{-1}$ for the normalized values of the parts of the refractive index that depends on the transition being investigated. Thus, the absorption coefficients are normalized to unity at resonance in zero fields. The Doppler widths correspond to the experimental values for the transitions, and the homogeneous widths were estimated to be the sum of the level widths. The two radiative intensities are characterized by the parameters $S_1=|G_1|^2/\Gamma_1\Gamma_{gm}$ and $S_4=|G_4|^2/\Gamma_4\Gamma_{gm}$.

Figure 2 depicts variations in the shape of the absorption and refraction spectra due to population motion and coherent effects, for neon atoms with zero projection of the velocity in the propagation direction of the radiation. We see that at zero offset of the strong radiation and selected values of the intensity of that radiation, with a ratio $N_l{:}N_g{:}N_m=100{:}50{:}85$ in the initial level populations, the absorption coefficient at the center of the line remains close to zero due to the splitting effect, even when the intensity of the auxiliary radiation is increased still further. At finite offset, the spectrum becomes asymmetric, and over a certain frequency range the radiation is amplified due to nonlinear interference effects. If the intensity and frequency offset of the strong field are increased still further, the amplification shifts to the region corresponding to backward Raman scattering, since for the unperturbed populations the two-photon transition is inverted. Under resonant conditions and at the selected intensities, there is no population inversion either in one-photon transitions or in two-photon transitions. Calculations yield $r_l{:}r_g{:}r_m=95.65{:}93.1{:}85$ for curve *1*, 95.46:95.35:85 for curve *2*, 96.3:87:85 for curve *3*, and 95.5:94.9:85 for curve *4*. The refractive index also undergoes a considerable change in the presence of the auxiliary radiation.

Because of Doppler shifts in a gas, the various atoms have differing offset ratios, so that the shape of nonlinear resonances is nearly symmetric, while their position is determined by the offset of the strong field within the Doppler profile and the relationship between the directions of propagation of the wave. Figure 3a shows the onset of inversionless amplification involving a short-wave Doppler-broadened transition in neon, and variation of the amplification with increasing intensity of the amplified radiation for the same initial level populations as in Fig. 2. Comparison with Fig. 2 suggests that because of atomic motion, the macroscopic coherence and the nonlinear interference effect do not disap-

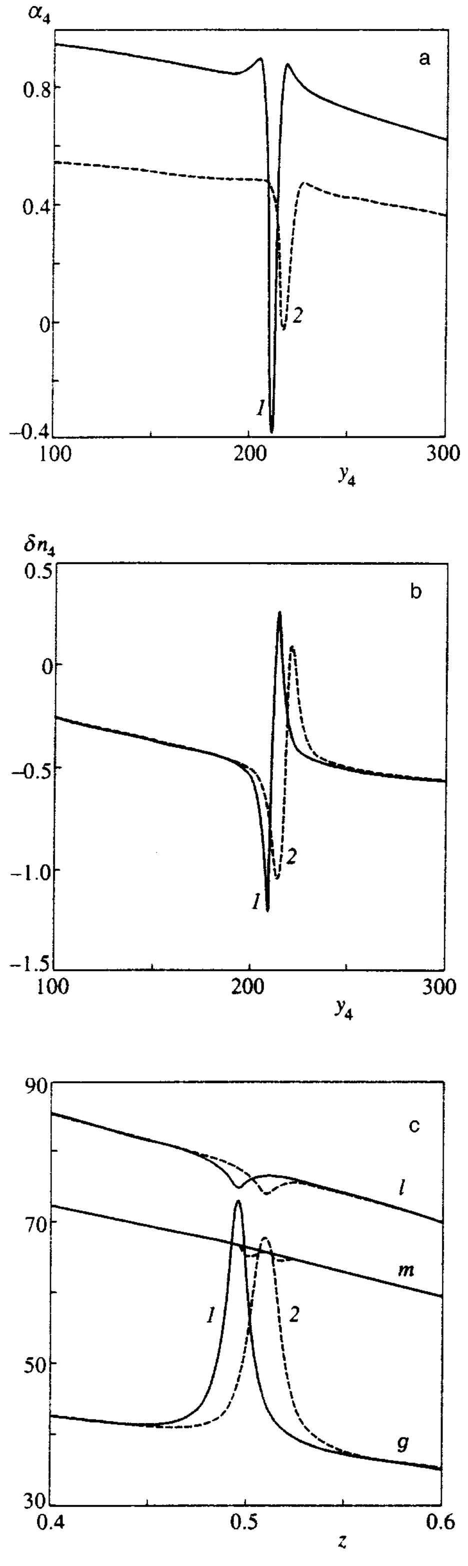


FIG. 3. Dependence of the velocity-averaged absorption coefficient (a) and refractive index (b) in the inhomogeneously broadened transition $l-m$ (l is the excited state, and the waves propagate in the same direction) in the presence of a strong field in the adjacent transition $l-g$ on the normalized offset y_4: curves *1*—$S_1=5$, $S_4=0$, and $\Omega_1=350\Gamma_{lg}$; curves *2*—$S_1=5$, $S_4=1$, and $\Omega_1=360\Gamma_{lg}$. c) Population distribution over the velocities $z=v/\overline{v}$ (the radiation frequencies correspond to the absorption minima). The ratios of the initial velocity-integrated unsaturated populations are the same as in Fig. 2: $N_l:N_g:N_m=100:50:85$. The solid curves correspond to curves *1* in Figs. 3a and b and the dashed curves, to the curves *2* in Figs. 3a and b.

pear. The relative variation in the refractive index and absorption coefficient proves to be even greater in Fig. 2. Figure 3b depicts the corresponding dispersion of the refractive index. Finally, Fig. 3c shows that there is no inversion of saturated populations in any of the transitions, and that the populations of the levels involved in the transition corresponding to the auxiliary radiation even out, which is also the case for the two-photon transition as the intensity of the amplified radiation increases (coherent population trapping). In the latter case, the velocity distribution of the population of the upper level m becomes slightly modulated, due to nonlinear optical effects in two strong fields.

Nonlinear interference effects in Doppler-broadened transitions have a pronounced angular anisotropy. For the V- and Λ-schemes with counterpropagating waves, the velocity packets of atoms in one- and two-photon resonances do not intersect, the velocity-averaged interference contribution vanishes, and the radiative interaction is governed solely by ladder transitions. Figure 4 shows that a resonance emerges on the opposite side of the Doppler profile, and amplification disappears despite the similar saturated populations at resonant velocities in comparison to the case depicted in Fig. 3. The experiments described in Refs. 10 and 11 investigated the special features of the generation of three-level gas lasers, due to nonlinear interference effects, where the lasing threshold exists only for waves propagating in the same direction. The theory of such lasers was examined in Refs. 4 and 13, including the case of unidirectional lasing with allowance for saturation in correlated transitions.[4,19]

The interference of contributions to the macroscopic polarization from atoms moving with different velocities manifests itself differently for short-wave and long-wave transitions. This is due to the reversal of the sign of the Doppler shift $(k_4-k_1)v$ for the two-photon resonance when the frequency ratio changes. After averaging over velocities, the splitting of atomic responses shows up only in long-wave transitions.[11,20]

The corresponding variations in the characteristics of inversionless amplification are shown in Fig. 5a, with the initial level populations changing in such a way that the unsaturated population differences between the amplified and auxiliary fields remain the same as in Fig. 4. Figure 5b depicts the corresponding velocity distribution of the level populations. Here the frequency of the probe radiation is tuned to the center of the dip in Fig. 5a (provided, of course, that the intensity of this radiation is nonzero). We see that as the intensity of the amplified radiation grows, the population of the upper resonant level m changes little, while the amplification and line profile change considerably.

Figures 6a–c show how strongly inversionless amplification depends on the ratio of the initial level populations. In accordance with the above analysis, if the saturated populations do not satisfy the above criteria, amplification is impossible no matter how strong (or weak) the fields are (Figs. 6a and b). In turn, the ratio of the saturated population differences depends on the initial populations, the relaxation characteristics of the transitions, and the intensities of the fields. If the above criteria are met, inversionless amplification becomes possible, and its value depends on the ratio of inten-

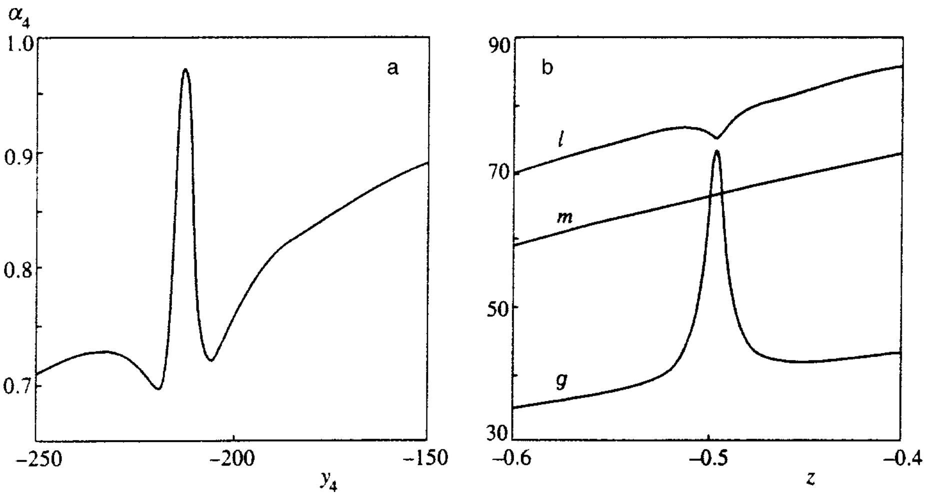


FIG. 4. The same as in Figs. 3a and c (curves *1*), but the waves propagate in opposite directions.

sities of the interacting light fields (Fig. 6c). By varying the population distribution by incoherent pumping, inversionless amplification can be maintained at a given level. This dependence determines the characteristics of lasing in an inversionless laser.

As noted earlier, open and closed systems of transitions have distinctive features, which are related both to any relationship among the relaxation parameters and to the interrelation between the incoherent pumping rate and the population of the lower level. Figures 7 and 8 illustrate some of these features using the sodium atom as a model, with the following relaxation characteristics of the $4P_{1/2}-3S$ ($\lambda_4=330\,\mathrm{nm}$) and $3P_{1/2}-3S$ ($\lambda_1=590\,\mathrm{nm}$) transitions: $\Gamma_m=9\times10^7\ \mathrm{s}^{-1}$ and $\Gamma_g=63\times10^7\ \mathrm{s}^{-1}$. By using weak probe radiation, Wang and Gao[21] analyzed the features of inversionless amplification involving the D_1 and D_2 transitions in sodium with allowance for Doppler broadening of the transitions and the hyperfine splitting of the lower level. We used the characteristics of a shorter-wave transition as a model, differing from the one discused in Ref. 21, in order to illustrate the dependence of inversionless amplification involving closed Doppler-broadened transitions on the intensity of the amplified radiation. As noted earlier, the motion of population differences and the splitting of resonances compete with nonlinear interference effects in the production of inversionless amplification. Hence, it is important to show that conditions close to optimal can be realized for a relatively broad class of objects.

Figure 7b (curve *1*) shows that an important feature of the present case is the dependence of the intensity of the upper level *m* on the intensity of the auxiliary radiation, even at vanishing intensities of the probe radiation. As noted earlier, this is due to the variation of the rate of incoherent pumping to this level as the population of the lower level changes. Curve *1* in figures 7a and b corresponds to the case in which 36% of the atoms are excited via incoherent pumping from the ground state to the level *m*, which for appropriate atomic concentrations can correspond to strong absorption of the probe radiation. If the system is irradiated by auxiliary light in resonance with the adjacent transition, weak short-wave radiation propagating in the same direction is amplified (at a level of approximately 50% of the initial absorption). Amplification takes place in the absence of inversion of saturated populations for all transitions. Curve *2* in Fig. 7a shows that amplification can vary considerably as the intensity of the amplified radiation increases, which is accompanied in the present scheme by significant variation in the populations of levels *m* and *l*. Figure 7c depicts the

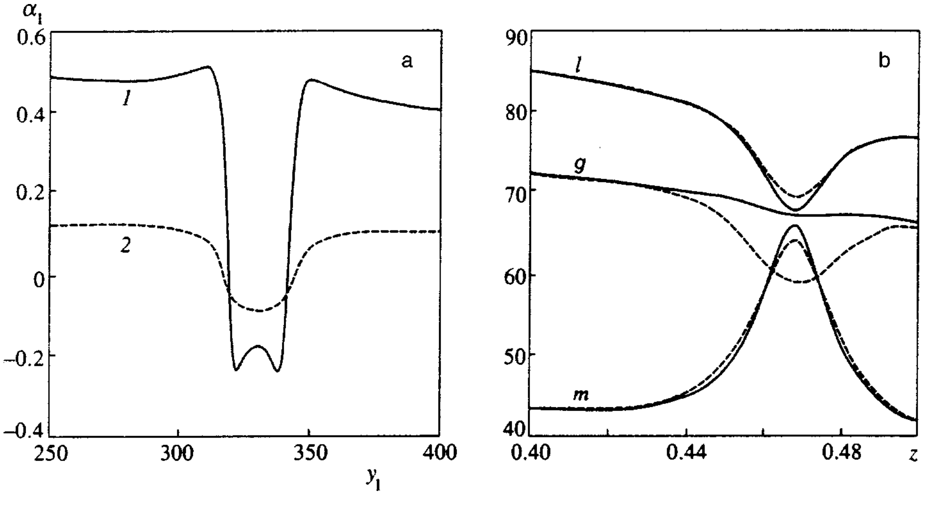


FIG. 5. a) Dependence of the velocity-averaged absorption coefficients in the inhomogeneously broadened long-wave transition $l-g$ on the normalized offset y_4; b) population distribution over the velocities $z=v/\bar{v}$ (l is the excited state, and the waves propagate in the same direction), in the presence of a strong field at the adjacent transition $l-m$ ($N_l{:}N_g{:}N_m=10{:}85{:}50$): curves *1*—$S_1=0.5$, $S_4=10$, and $\Omega_4=200\Gamma_{lm}$; curves *2*—$S_1=10$, $S_4=10$, and $\Omega_4=200\Gamma_{lm}$. In Fig. 5b the radiation frequencies correspond to the center of the dip (the solid curves correspond to curves *1* in Fig. 5a, and the dashed curves, to curves *2* in Fig. 5a).

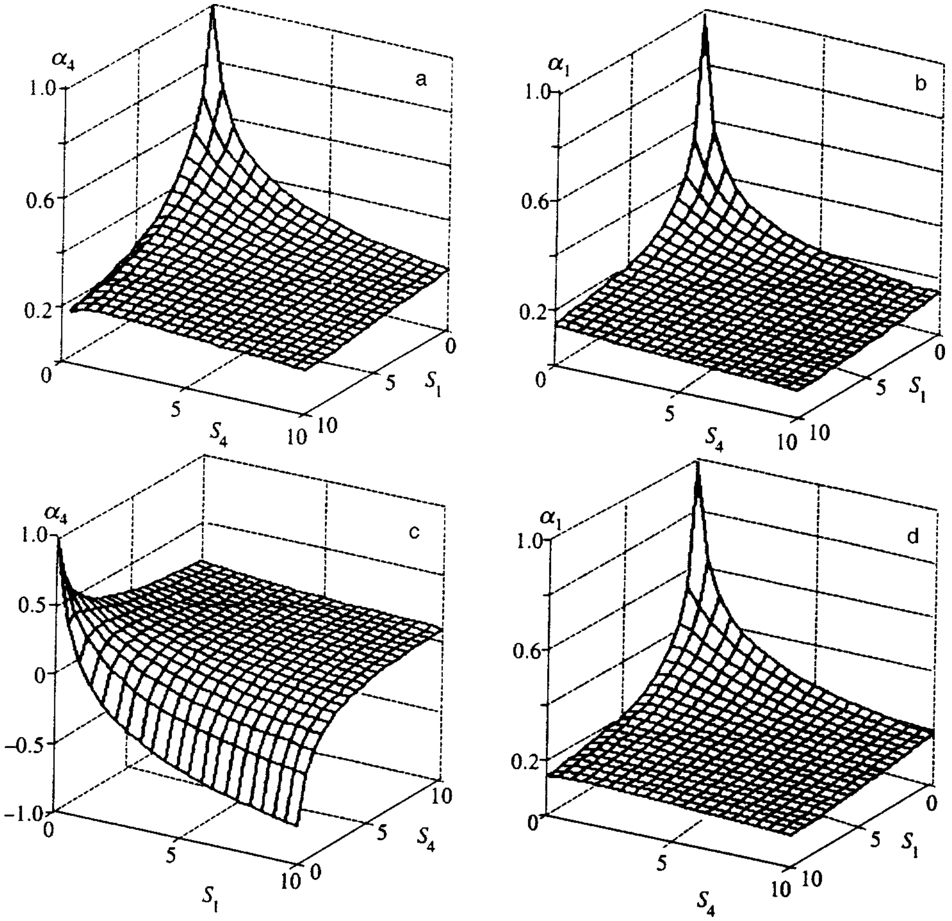


FIG. 6. Dependence of the velocity-averaged values of the absorption coefficients α_1 and α_4 at zero offset of either field (l is the excited state, and the waves propagate in the same direction) on the normalized radiative intensities S_1 and S_2: (a) and (b), $N_l{:}N_g{:}N_m = 100{:}50{:}70$; (c) and (d), $N_l{:}N_g{:}N_m = 100{:}50{:}85$.

velocity and energy-level distribution of atoms that corresponds to the disappearance of both absorption and amplification. It is interesting to note how this distribution changes abruptly due to nonlinear optical effects involving an increase in the intensity of the weaker radiation.

Figure 8 shows that by varying the initial distribution via incoherent pumping and the intensity of the auxiliary radiation, we can maintain, over a certain range of values, a constant level of amplification as the intensity of the amplified radiation rises.

5. CONCLUSION

Our results contribute to the development of the theory of resonant nonlinear interference processes involving quantum Doppler-broadened transitions in strong electromagnetic fields for open and closed three-level configurations, with allowance for incoherent pumping of levels and for processes accompanying the coherent interaction of light fields.

The expressions derived in this paper make it possible to analyze and compare, from a consistent standpoint, the manifestations of quantum interference processes in various V-, Λ-, and ladder configurations of interacting transitions, with allowance for processes that strongly affect the choice and optimization of the conditions of experiments in the design of lasers that do not require population inversion. These capabilities have been demonstrated through the use of numerical models characteristic of open and closed configurations of optical transitions.

The authors are grateful to Prof. L. J. F. Hermans of the University of Leiden and The Netherlands Scientific Society (NWO) for support of the present work. This work was supported by grants from the Russian Fund for Fundamental Research (Grants No. 96-02-00010C, No. 96-02-00016G, and No. 97-02-16092).

APPENDICES

Here we give the outline and results of calculations of the refractive index, the absorption coefficient, and level populations for various open and closed configurations of energy levels.

A. V-SCHEME (FIELDS E_1 AND E_4)

A.1. Open configuration

The equations for the density matrix are

$$L_{gm}\rho_{gm} = i(\rho_{gl}V_{lm} - V_{gl}\rho_{lm})$$

$$= L_{41}\rho_{41} = i(\rho_1^* V_4 - V_1^* \rho_4),$$

$$L_1\rho_1 = -i\{V_1(\rho_g - \rho_l) + \rho_{41}^* V_4\},$$

$$L_4\rho_4 = -i\{V_4(\rho_m - \rho_l) + \rho_{41}V_1\},$$

$$L_g\rho_g = -2\,\mathrm{Re}(iV_1^*\rho_1) + q_g,$$

$$L_m\rho_m = -2\,\mathrm{Re}(iV_4^*\rho_4) + q_m,$$

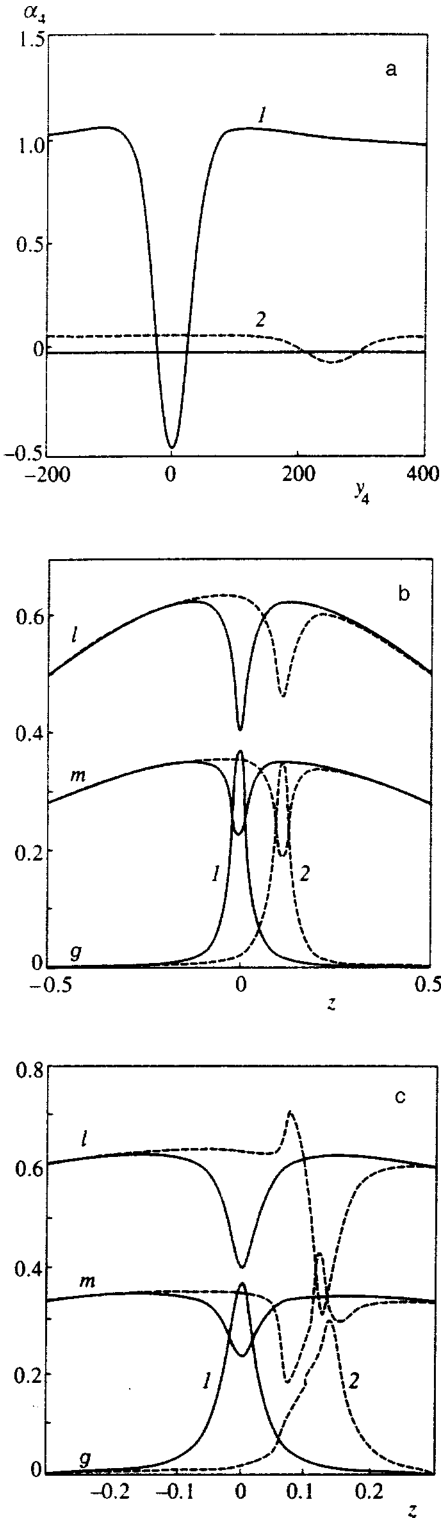


FIG. 7. a) Dependence on the normalized offset y_4 of the velocity-averaged absorption coefficient in the inhomogeneously broadened long-wave transition $l-m$ (l is the ground state, and the waves propagate in the same direction) in the presence of a strong field at the adjacent transition $l-g$: curves *1*—$S_1=10$, $S_4=0$, and $y_1=0$; curves *2*—$S_1=10$, $S_4=20$, and $y_4=20$. b,c) Saturated population distributions over the velocities $z=v/\bar{v}$: in (b), the radiation frequencies correspond to the absorption minima; in (c), the radiation frequencies correspond to the right (*1*) and left (*2*) zero-absorption points of the corresponding curves. $N_l:N_g:N_m=64:0:36$.

$$L_l\rho_l=q_l-2\ \mathrm{Re}(iV_1\rho_1^*+iV_4\rho_4^*)+\gamma_4\rho_m+\gamma_1\rho_g. \quad \text{(A1)}$$

For the amplitudes of the off-diagonal elements of the density matrix (r_1, r_{41}, and r_4), in the stationary case we have a system of algebraic equations, whose solution is

$$r_1=i\,\frac{G_1}{P_1F^*}\,[\Delta r_1(1+g_{1v}^*)-u_{1v}^*\Delta r_4],$$

$$r_{41}=\frac{G_1^*G_4}{P_{41}P_4P_1^*F}\,[\Delta r_1P_4+\Delta r_4P_1^*],$$

$$r_4=i\,\frac{G_4}{P_4F}\,[\Delta r_4(1+u_{2v})-\Delta r_1g_{2v}],$$

$$u_{1v}=\frac{|G_4|^2}{P_4P_{41}},\quad u_{2v}=\frac{|G_4|^2}{P_1^*P_{41}},$$

$$g_{1v}=\frac{|G_1|^2}{P_4P_{41}},\quad g_{2v}=\frac{|G_1|^2}{P_1^*P_{41}},$$

$$P_1=P_{lg}=\Gamma_1+i\Omega_1,\quad P_{41}=P_{lg}=\Gamma_{41}+i(\Omega_4-\Omega_1),$$

$$F=1+g_{1v}+u_{2v},\quad \Delta r_1=r_l-r_g,\quad \Omega_1=\omega_1-\omega_{gl}. \quad \text{(A2)}$$

Plugging (A2) into the equations for the diagonal elements of the density matrix in (A1), we find that

$$r_m=n_m+(1-a_{44v})\kappa_4[\Delta r_4F_{1v}-\Delta r_1F_{2v}],$$

$$r_l=n_l-\Delta r_4(a_{44v}\kappa_4F_{1v}-a_{31v}\kappa_1F_{4v})$$

$$+\Delta r_1(a_{44v}\kappa_4F_{2v}-a_{31v}\kappa_1F_{3v}), \quad \text{(A3)}$$

$$r_g=n_g-(1-a_{31v})\kappa_1[\Delta r_4F_{4v}-\Delta r_1F_{3v}],$$

$$a_{31v}=\frac{\Gamma_g-\gamma_1}{\Gamma_g+\Gamma_l-\gamma_1},\quad a_{44v}=\frac{\Gamma_m-\gamma_4}{\Gamma_m+\Gamma_l-\gamma_4};$$

$$F_{1v}+if_{1v}=\frac{\Gamma_4(1+u_{2v})}{P_4F},\quad F_{2v}+if_{2v}=\frac{\Gamma_4g_{2v}}{P_4F},$$

$$F_{3v}+if_{3v}=\frac{\Gamma_1(1+g_{1v})}{P_1^*F},\quad F_{4v}+if_{4v}=\frac{\Gamma_1u_{1v}}{P_1^*F}. \quad \text{(A4)}$$

The saturation parameters for the first and fourth transitions are

$$\kappa_1=\kappa_{lg}=\frac{2|G_1|^2(\Gamma_l+\Gamma_g-\gamma_1)}{\Gamma_l\Gamma_g\Gamma_1},$$

$$\kappa_4=\frac{2|G_4|^2(\Gamma_l+\Gamma_m-\gamma_4)}{\Gamma_l\Gamma_m\Gamma_4}. \quad \text{(A5)}$$

Calculation of the population difference for the corresponding transitions yields

$$\Delta r_4=\frac{\Delta n_4X_{2v}-\Delta n_1X_{3v}}{X_{1v}X_{2v}-X_{3v}X_{4v}},\quad \Delta r_1=\frac{\Delta n_1X_{1v}-\Delta n_4X_{4v}}{X_{1v}X_{2v}-X_{3v}X_{4v}},$$

$$X_{1v}=1-a_{31v}\kappa_1F_{4v}+\kappa_4F_{1v},$$

$$X_{2v}=1-a_{44v}\kappa_4F_{2v}+\kappa_1F_{3v},$$

$$X_{3v}=a_{31v}\kappa_1F_{3v}-\kappa_4F_{2v},$$

$$X_{4v}=a_{44v}\kappa_4F_{1v}-\kappa_1F_{4v}. \quad \text{(A6)}$$

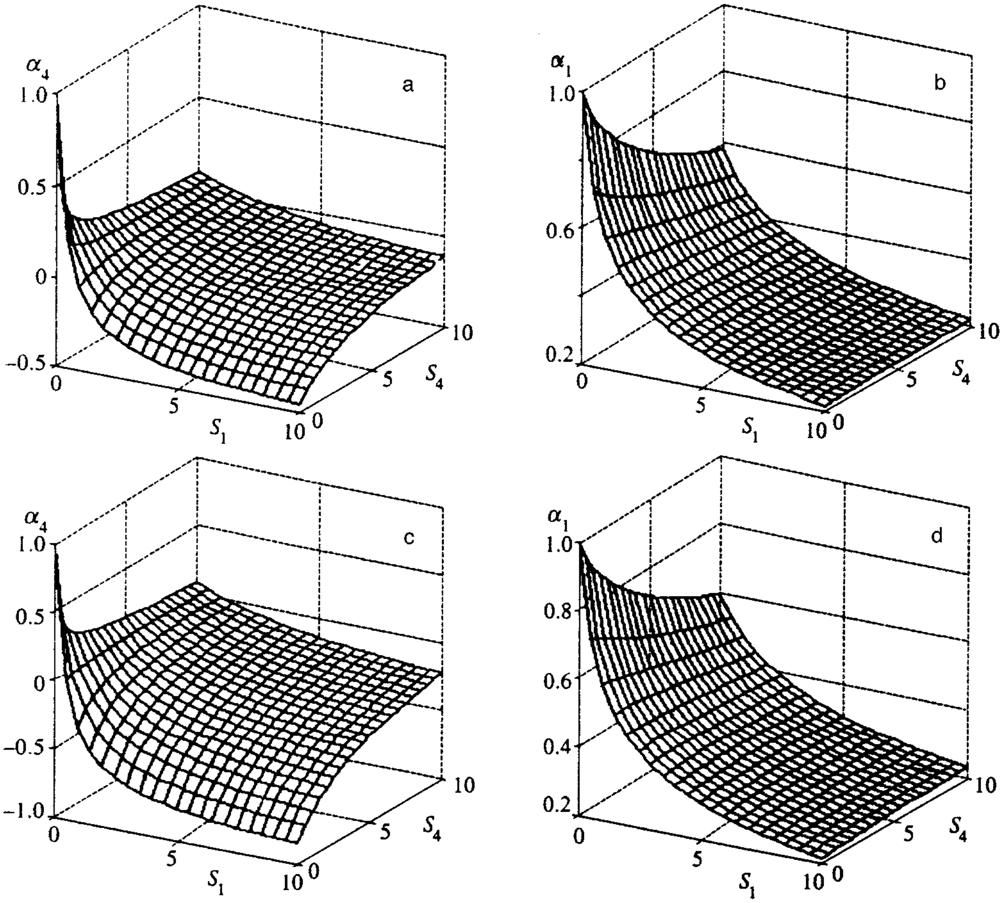


FIG. 8. Dependence of the velocity-averaged absorption coefficients α_1 and α_4 on the normalized radiative intensities S_1 and S_2 at zero offset of both fields (l is the ground state, and the waves propagate in the same direction): a,b), N_l:N_g:N_m=64:0:36; c,d), N_l:N_g:N_m=60:0:40.

Using these solutions, we obtain

$$\frac{\alpha_4}{\alpha_{04}}=F_{1v}\frac{\Delta r_4}{\Delta n_4}-F_{2v}\frac{\Delta r_1}{\Delta n_4},\quad \frac{\alpha_1}{\alpha_{01}}=F_{3v}\frac{\Delta r_1}{\Delta n_1}-F_{4v}\frac{\Delta r_4}{\Delta n_1}, \tag{A7}$$

$$\frac{n_4(\Omega_4)-n_4^{\mathrm{nr}}}{2(n_{4\,\mathrm{max}}^0-n_4^{\mathrm{nr}})}=f_{1v}\frac{\Delta r_4}{\Delta n_4}-f_{2v}\frac{\Delta r_1}{\Delta n_4},$$

$$-\frac{n_1(\Omega_1)-n_1^{\mathrm{nr}}}{2(n_{1\,\mathrm{max}}^0-n_1^{\mathrm{nr}})}=f_{3v}\frac{\Delta r_1}{\Delta n_1}-f_{4v}\frac{\Delta r_4}{\Delta n_1}. \tag{A8}$$

A.2. Closed configuration

As noted earlier, in this case the incoherent excitation rates depend on induced transitions in the system, as reflected in the equation for the populations:

$$L_g\rho_g=w_g\rho_l-2\,\mathrm{Re}(iV_1^*\rho_1),$$

$$L_m\rho_m=w_m\rho_l-2\,\mathrm{Re}(iV_4^*\rho_4),$$

$$\rho_l=1-\rho_m-\rho_g\,. \tag{A9}$$

Here w_g and w_m are the probabilities of arrival at the given level.

Since the form of the equations for the off-diagonal elements in a closed system is the same as before, the corresponding stationary solutions for the population differences, the absorption and amplification coefficients, and the refractive index can be obtained from (A6) and (A7) via the simple substitutions

$$X_{1v}=1+0.5(1+\Delta n_4)\kappa_4F_{1v}-0.5\Delta n_4\kappa_1F_{4v}\,,$$

$$X_{2v}=1+0.5(1+\Delta n_1)\kappa_1F_{3v}-0.5\Delta n_1\kappa_4F_{2v}\,,$$

$$X_{3v}=0.5\Delta n_4\kappa_1F_{3v}-0.5(1+\Delta n_4)\kappa_4F_{2v}\,,$$

$$X_{4v}=0.5\Delta n_1\kappa_4F_{1v}-0.5(1+\Delta n_1)\kappa_1F_{4v}\,,$$

$$\kappa_1=\frac{4|G_1|^2}{\Gamma_g\Gamma_1},\quad \kappa_4=\frac{4|G_4|^2}{\Gamma_m\Gamma_4}.$$

The initial and saturated level populations are

$$n_l=\frac{1}{1+w_g/\Gamma_g+w_m/\Gamma_m},$$

$$n_g=\frac{w_g/\Gamma_g}{1+w_g/\Gamma_g+w_m/\Gamma_m},$$

$$n_m=\frac{w_m/\Gamma_m}{1+w_g/\Gamma_g+w_m/\Gamma_m},$$

$$r_m=\tfrac{1}{3}[1+\Delta r_1-2\Delta r_4],\quad r_g=\tfrac{1}{3}[1+\Delta r_4-2\Delta r_1],$$

$$r_l=\tfrac{1}{3}[1+\Delta r_1+\Delta r_4].$$

B. Λ-SCHEME (FIELDS E_3 AND E_4)

B.1. Open configuration

In this case the initial system of equations assumes the form

$$\begin{aligned}
L_{ln}\rho_{ln}&=i(\rho_{lm}V_{mn}-V_{lm}\rho_{mn})\\
&=L_{43}\rho_{43}=i(\rho_4V_3^*-V_4\rho_3^*),\\
L_4\rho_4&=-i[V_4(\rho_m-\rho_l)-\rho_{43}V_3],\\
L_3\rho_3&=-i[V_3(\rho_m-\rho_n)+\rho_{43}^*V_4],\\
L_n\rho_n&=-2\,\mathrm{Re}(iV_3\rho_3^*)+\gamma_3\rho_m+q_n,\\
L_m\rho_m&=-2\,\mathrm{Re}(iV_3^*\rho_3+iV_4^*\rho_4)+q_m,\\
L_l\rho_l&=q_l-2\,\mathrm{Re}(iV_4\rho_4^*)+\gamma_4\rho_m.
\end{aligned}\tag{B1}$$

As in Appendix A, in the stationary case we can reduce this system to algebraic form. The solution is

$$\begin{aligned}
r_3&=i\,\frac{G_3}{P_3F^*}\,[\Delta r_3(1+g_{1l}^*)-u_{1l}^*\Delta r_4],\\
r_{43}&=-\frac{G_3^*G_4}{P_{43}P_4P_3^*F}\,[\Delta r_3P_4+\Delta r_4P_3^*],\\
r_4&=i\,\frac{G_4}{P_4F}\,[\Delta r_4(1+u_{2l})-\Delta r_3g_{2l}],\\
u_{1l}&=\frac{|G_4|^2}{P_4P_{43}},\quad u_{2l}=\frac{|G_4|^2}{P_3^*P_{43}},
\end{aligned}\tag{B2}$$

$$\begin{aligned}
g_{1l}&=\frac{|G_3|^2}{P_4P_{43}},\quad g_{2l}=\frac{|G_3|^2}{P_3^*P_{43}},\\
P_{nm}&=P_3=\Gamma_3+i\Omega_3,\quad P_{ln}=P_{43}=\Gamma_{43}+i(\Omega_4-\Omega_3),\\
F&=1+g_{1l}+u_{2l},\quad \Delta r_3=r_n-r_m,\quad \Omega_3=\omega_3-\omega_{mn};\\
r_l&=n_l+\Delta r_3[c_2\kappa_4F_{2l}+b_1\kappa_3F_{3l}]\\
&\quad-\Delta r_4[c_2\kappa_4F_{1l}+b_1\kappa_3F_{4l}],\\
r_n&=n_n-\Delta r_3(c_1\kappa_4F_{2l}+b_2\kappa_3F_{3l})\\
&\quad+\Delta r_4(c_1\kappa_4F_{1l}+b_2\kappa_3F_{4l}),
\end{aligned}\tag{B3}$$

$$\begin{aligned}
r_m&=n_m-\Delta r_3[\kappa_4F_{2l}(1-c_2)-\kappa_3F_{3l}(1-b_2)]\\
&\quad+\Delta r_4[\kappa_4F_{1l}(1-c_2)-(1-b_2)\kappa_3F_{4l}],\\
b_1&=\frac{\Gamma_n}{\Gamma_l}\,\frac{\gamma_4}{\Gamma_m+\Gamma_n-\gamma_3},\quad b_2=\frac{\Gamma_m-\gamma_3}{\Gamma_m+\Gamma_n-\gamma_3},\\
c_1&=\frac{\Gamma_l}{\Gamma_n}\,\frac{\gamma_3}{\Gamma_l+\Gamma_m-\gamma_4},\quad c_2=\frac{\Gamma_m-\gamma_4}{\Gamma_l+\Gamma_m-\gamma_4};\\
\kappa_3&=\frac{2|G_3|^2(\Gamma_m+\Gamma_n-\gamma_3)}{\Gamma_m\Gamma_n\Gamma_3},
\end{aligned}\tag{B4}$$

$$\kappa_4=\frac{2|G_4|^2(\Gamma_m+\Gamma_l-\gamma_4)}{\Gamma_m\Gamma_l\Gamma_4};$$

$$F_{1l}+if_{1l}=\frac{\Gamma_4(1+u_{2l})}{P_4F},\quad F_{2l}+if_{2l}=\frac{\Gamma_4g_{2l}}{P_4F},\tag{B5}$$

$$\begin{aligned}
F_{3l}+if_{3l}&=\frac{\Gamma_3(1+g_{1l})}{P_3^*F},\quad F_{4l}+if_{4l}=\frac{\Gamma_3u_{1l}}{P_3^*F};\\
\Delta r_4&=\frac{\Delta n_4X_{2l}-\Delta n_3X_{3l}}{X_{1l}X_{2l}-X_{3l}X_{4l}},\quad \Delta r_3=\frac{\Delta n_3X_{1l}-\Delta n_4X_{4l}}{X_{1l}X_{2l}-X_{3l}X_{4l}},\\
X_{1l}&=1+\kappa_4F_{1l}-(1-b_1-b_2)\kappa_3F_{4l},\\
X_{2l}&=1+\kappa_3F_{3l}-(1-c_1-c_2)\kappa_4F_{2l},\\
X_{3l}&=(1-b_1-b_2)\kappa_3F_{3l}-\kappa_4F_{2l},\\
X_{4l}&=(1-c_1-c_2)\kappa_4F_{1l}-\kappa_3F_{4l}.
\end{aligned}\tag{B6}$$

The expressions for the amplification and absorption coefficients and the refractive index for the $m-l$ and $m-n$ transitions are

$$\frac{\alpha_4}{\alpha_{04}}=F_{1l}\frac{\Delta r_4}{\Delta n_4}-F_{2l}\frac{\Delta r_3}{\Delta n_4},\quad \frac{\alpha_3}{\alpha_{03}}=F_{3l}\frac{\Delta r_3}{\Delta n_3}-F_{4l}\frac{\Delta r_4}{\Delta n_3},\tag{B7}$$

$$\begin{aligned}
&\frac{n_4(\Omega_4)-n_4^{\mathrm{nr}}}{2(n_{4\,\mathrm{max}}^0-n_4^{\mathrm{nr}})}=f_{1l}\frac{\Delta r_4}{\Delta n_4}-f_{2l}\frac{\Delta r_1}{\Delta n_4},\\
&-\frac{n_3(\Omega_3)-n_3^{\mathrm{nr}}}{2(n_{3\,\mathrm{max}}^0-n_3^{\mathrm{nr}})}=f_{3l}\frac{\Delta r_1}{\Delta n_1}-f_{4l}\frac{\Delta r_4}{\Delta n_1}.
\end{aligned}\tag{B8}$$

B.2. Closed configuration

The initial system of equations is

$$\begin{aligned}
L_n\rho_n&=w_n\rho_l-2\,\mathrm{Re}(iV_3\rho_3^*)+\gamma_3\rho_m,\\
L_m\rho_m&=w_m\rho_l-2\,\mathrm{Re}(iV_3^*\rho_3+iV_4^*\rho_4),\\
\rho_l&=1-\rho_n-\rho_m.
\end{aligned}\tag{B9}$$

As in the previous section, we can represent the stationary solution of the algebraic system of equations in the form (B6), where

$$\begin{aligned}
X_{1l}&=1+0.5\left[1+\Delta n_4\left(1+\frac{\gamma_3}{\Gamma_n}\right)\right]\kappa_4F_{1l}\\
&\quad-[1-b_2+\Delta n_4(1-2b_2)]\kappa_3F_{4l},\\
X_{2l}&=1+[1+\Delta n_3(1-2b_2)]\kappa_3F_{3l}\\
&\quad-0.5\left[1-\frac{\gamma_3}{\Gamma_n}+\Delta n_3\left(1+\frac{\gamma_3}{\Gamma_n}\right)\right]\kappa_4F_{2l},\\
X_{3l}&=[1-b_2+\Delta n_4(1-2b_2)]\kappa_3F_{3l}\\
&\quad-0.5\left[1+\Delta n_4\left(1+\frac{\gamma_3}{\Gamma_n}\right)\right]\kappa_4F_{2l},\\
X_{4l}&=0.5\left[1-\frac{\gamma_3}{\Gamma_n}+\Delta n_3\left(1+\frac{\gamma_3}{\Gamma_n}\right)\right]\kappa_4F_{1l}\\
&\quad-[1+\Delta n_3(1-2b_2)]\kappa_3F_{4l},
\end{aligned}$$

$$\kappa_4=\frac{4|G_4|^2}{\Gamma_m\Gamma_4},$$

$$n_l=\frac{1}{1+w_n'/\Gamma_n+w_m/\Gamma_m},\quad n_n=\frac{w_n'/\Gamma_n}{1+w_n'/\Gamma_n+w_m/\Gamma_m},$$

$$n_m=\frac{w_m/\Gamma_m}{1+w_n'/\Gamma_n+w_m/\Gamma_m},\quad w_n'=w_n+\frac{\gamma_3}{\Gamma_n}\,w_m.$$

The saturated level populations can be expressed in terms of the solutions (B6) for the population differences ($\Delta r_3=r_n-r_m$ and $\Delta r_4=r_l-r_m$)

$$r_m=\tfrac{1}{3}[1-\Delta r_3-\Delta r_4],\quad r_n=\tfrac{1}{3}[1+2\Delta r_3-\Delta r_4],$$
$$r_l=\tfrac{1}{3}[1+2\Delta r_4-\Delta r_3].$$

C. H-SCHEME (FIELDS E_4 AND E_2)

C.1. Open configuration

The initial equations for the density matrix are

$$\begin{aligned}
&L_{lf}\rho_{lf}=i(\rho_{lm}V_{mf}-V_{lm}\rho_{mf})=L_{42}\rho_{42}=i(\rho_4V_2-V_4\rho_2),\\
&L_2\rho_2=-i[V_2(\rho_f-\rho_m)+\rho_{42}V_4^*],\\
&L_4\rho_4=-i[V_4(\rho_m-\rho_l)-\rho_{42}V_2^*],\\
&L_m\rho_m=q_m-2\,\mathrm{Re}(iV_4^*\rho_4+iV_2\rho_2^*)+\gamma_2\rho_f,\\
&L_f\rho_f=-2\,\mathrm{Re}(iV_2^*\rho_2)+q_f,\\
&L_l\rho_l=-2\,\mathrm{Re}(iV_4\rho_4^*)+\gamma_4\rho_m+q_l.
\end{aligned}\tag{C1}$$

The stationary solutions of the algebraic system of equations for the amplitudes of off-diagonal elements have a similar form:

$$\begin{aligned}
&r_4=i\,\frac{G_4}{P_4F}\,[\Delta r_4(1+u_{2h})+g_{2h}\Delta r_2],\\
&r_{42}=-\frac{G_4G_2}{P_{42}P_4P_2F}\,[\Delta r_2P_4-\Delta r_4P_2],\\
&r_2=i\,\frac{G_2}{P_2F}\,[\Delta r_2(1+g_{1h})+\Delta r_4u_{1h}],\\
&g_{1h}=\frac{|G_2|^2}{P_4P_{42}},\quad g_{2h}=\frac{|G_2|^2}{P_2P_{42}},\\
&u_{1h}=\frac{|G_4|^2}{P_4P_{42}},\quad u_{2h}=\frac{|G_4|^2}{P_2P_{42}},\\
&P_{mf}=P_2=\Gamma_2+i\Omega_2,\quad P_{lf}=P_{42}=\Gamma_{42}+i(\Omega_4+\Omega_2),\\
&F=1+g_{1h}+u_{2h},\quad \Delta r_2=r_m-r_f,\quad \Omega_2=\omega_2-\omega_{fm}.
\end{aligned}\tag{C2}$$

The equations for the populations and population differences can be written as follows:

$$\begin{aligned}
&r_f=n_f+(1-d_2)\kappa_2[\Delta r_4F_{4h}+\Delta r_2F_{3h}],\\
&r_l=n_l-\Delta r_4(\kappa_2d_1F_{4h}+\kappa_4b_2F_{1h})\\
&\qquad-\Delta r_2(\kappa_2d_1F_{3h}+\kappa_4b_2F_{2h}),
\end{aligned}\tag{C3}$$

$$\begin{aligned}
&r_m=n_m+\Delta r_4[\kappa_4F_{1h}(1-b_2)-\kappa_2F_{4h}d_2]\\
&\qquad+\Delta r_2[\kappa_4F_{2h}(1-b_2)-\kappa_2F_{3h}d_2],\\
&b_2=\frac{\Gamma_m-\gamma_4}{\Gamma_m+\Gamma_l-\gamma_4},\quad d_1=d_2\,\frac{\gamma_4}{\Gamma_l},\quad d_2=\frac{\Gamma_f-\gamma_2}{\Gamma_m+\Gamma_f-\gamma_2};\\
&F_{1h}+if_{1h}=\frac{\Gamma_4(1+u_{2h})}{P_4F},\quad F_{2h}+if_{2h}=\frac{\Gamma_4g_{2h}}{P_4F},\\
&F_{3h}+if_{3h}=\frac{\Gamma_2(1+g_{1h})}{P_2F},\quad F_{4h}+if_{4h}=\frac{\Gamma_2u_{1h}}{P_2F};
\end{aligned}\tag{C4}$$

$$\begin{aligned}
&\kappa_4=\frac{2|G_4|^2(\Gamma_l+\Gamma_m-\gamma_4)}{\Gamma_l\Gamma_m\Gamma_4},\\
&\kappa_2=\frac{2|G_2|^2(\Gamma_f+\Gamma_m-\gamma_2)}{\Gamma_f\Gamma_m\Gamma_2};
\end{aligned}\tag{C5}$$

$$\begin{aligned}
&\Delta r_4=\frac{\Delta n_4X_{2h}+\Delta n_2X_{3h}}{X_{1h}X_{2h}-X_{3h}X_{4h}},\\
&\Delta r_2=\frac{\Delta n_2X_{1h}+\Delta n_4X_{4h}}{X_{1h}X_{2h}-X_{3h}X_{4h}},
\end{aligned}\tag{C6}$$

Here $\Delta n_4=n_l-n_m$, $\Delta n_2=n_m-n_f$, and

$$\begin{aligned}
&X_{1l}=1-d_2\left(1-\frac{\gamma_4}{\Gamma_l}\right)\kappa_2F_{4h}+\kappa_4F_{1h},\\
&X_{2h}=1+\kappa_2F_{3h}-(1-b_2)\kappa_4F_{2h},\\
&X_{3h}=d_2\left(1-\frac{\gamma_4}{\Gamma_l}\right)\kappa_2F_{3h}-\kappa_4F_{2h},\\
&X_{4h}=(1-b_2)\kappa_4F_{1h}-\kappa_2F_{4h}.
\end{aligned}$$

The expressions for the amplification and absorption coefficients and the refractive index for the $m-l$ and $f-m$ transitions are

$$\begin{aligned}
&\frac{\alpha_4}{\alpha_{04}}=F_{1h}\frac{\Delta r_4}{\Delta n_4}+F_{2h}\frac{\Delta r_2}{\Delta n_4},\\
&\frac{\alpha_2}{\alpha_{02}}=F_{3h}\frac{\Delta r_2}{\Delta n_2}+F_{4h}\frac{\Delta r_4}{\Delta n_2},
\end{aligned}\tag{C7}$$

$$\begin{aligned}
&\frac{n_4(\Omega_4)-n_4^{\mathrm{nr}}}{2(n_{4\,\max}^0-n_4^{\mathrm{nr}})}=f_{1h}\frac{\Delta r_4}{\Delta n_4}+f_{2h}\frac{\Delta r_2}{\Delta n_4},\\
&\frac{n_2(\Omega_2)-n_2^{\mathrm{nr}}}{2(n_{2\,\max}^0-n_2^{\mathrm{nr}})}=f_{3h}\frac{\Delta r_2}{\Delta n_2}+f_{4h}\frac{\Delta r_4}{\Delta n_2}.
\end{aligned}\tag{C8}$$

C.2. Closed configuration

The equations for the populations are

$$\begin{aligned}
&L_m\rho_m=w_m\rho_l-2\,\mathrm{Re}(iV_4^*\rho_4+iV_2\rho_2^*)+\gamma_2\rho_f,\\
&L_f\rho_f=w_f\rho_l-2\,\mathrm{Re}(iV_2^*\rho_2),\quad \rho_l=1-\rho_m-\rho_f.
\end{aligned}\tag{C9}$$

Using the stationary solutions (C2) for the off-diagonal elements, we find that the transition from (C6) for the open configuration can be achieved via the substitutions

$$X_{1h}=1+0.5[1+\Delta n_4]\kappa_4F_{1h}$$

$$
\begin{aligned}
&\qquad -[d_2-\Delta n_4(1-2d_2)]\kappa_2 F_{4h},\\
&X_{2h}=1+[1+\Delta n_2(1-2d_2)]\kappa_2 F_{3h}\\
&\qquad -0.5[1-\Delta n_2]\kappa_4 F_{2h},\\
&X_{3h}=[d_2-\Delta n_4(1-2d_2)]\kappa_2 F_{3h}-0.5[1+\Delta n_4]\kappa_4 F_{2h},\\
&X_{4h}=0.5[1-\Delta n_2]\kappa_4 F_{1h}-[1+\Delta n_2(1-2d_2)]\kappa_2 F_{4h},\\
&\kappa_4=\frac{4|G_4|^2}{\Gamma_m\Gamma_4}.
\end{aligned}
$$

Here the formulas for the initial level populations are

$$
\begin{aligned}
&n_l=\frac{1}{1+w_f/\Gamma_f+w_m'/\Gamma_m},\\
&n_f=\frac{w_f/\Gamma_f}{1+w_f/\Gamma_f+w_m'/\Gamma_m},\\
&n_m=\frac{w_m'/\Gamma_m}{1+w_f/\Gamma_f+w_m'/\Gamma_m},\quad w_m'=w_m+\frac{\gamma_2}{\Gamma_m}w_f.
\end{aligned}
$$

The saturated level populations can be expressed in terms of the population differences ($\Delta r_2=r_m-r_f$ and $\Delta r_4=r_l-r_m$) by using the equation $r_l=1-r_f-r_m$:

$$
\begin{aligned}
&r_m=\tfrac{1}{3}[1+\Delta r_2-\Delta r_4],\quad r_f=\tfrac{1}{3}[1-\Delta r_4-2\Delta r_2],\\
&r_l=\tfrac{1}{3}[1+\Delta r_2+2\Delta r_4].
\end{aligned}
$$

*)Fax: (3912)438923; e-mail: popov@cc.krascience.rssi.ru

---

[1] M. P. Chaĭka, *Interference of Degenerate Atomic States* [in Russian], Leningrad Univ. Press, Leningrad (1975); E. B. Aleksandrov, G. I. Khvostenko, and M. P. Chaĭka, *Interference of Atomic States*, Springer-Verlag, New York (1993).

[2] P. A. Apanasevich, *Basics of the Interaction of Light with Matter* [in Russian], Nauka i Tekhnika, Minsk (1977).

[3] S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in the Spectra of Atoms and Molecules* [in Russian], Nauka, Novosibirsk (1979); S. G. Rautian and A. M. Shalagin, *Kinetic Problems of Nonlinear Spectroscopy*, North-Holland, Amsterdam (1991).

[4] A. K. Popov, *Introduction to Nonlinear Spectroscopy* [in Russian], Novosibirsk, Nauka (1983).

[5] A. Javan, Phys. Rev. **107**, 1579 (1957); V. M. Kontorovich and A. M. Prokhorov, Zh. Éksp. Teor. Fiz. **33**, 1428 (1957) [Sov. Phys. JETP **6**, 1100 (1958)]; T. Yajima and K. Shimoda, J. Phys. Soc. Jpn. **15**, 1668 (1960); Adv. Quantum Electron. **3**, 548 (1961); V. M. Faĭn and Ya. I. Khanin, *Quantum Radiophysics*, Vol. 1: V. M. Faĭn, *Photons and Nonlinear Media* [in Russian], Sovetskoe Radio, Moscow (1972), p. 389 [English translation: *Quantum Electronics*, 2 vols., Pergamon Press, Oxford (1969)].

[6] D. N. Klyshko, Yu. S. Konstantinov, and V. S. Tumanov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **8**, 513 (1965); A. M. Bonch-Bruevich, V. A. Khodovoĭ, and N. A. Chigir', Usp. Fiz. Nauk **93**, 71 (1967); Zh. Éksp. Teor. Fiz. **67**, 2069 (1974) [Sov. Phys. JETP **40**, 1027 (1975)]; A. M. Bonch-Bruevich, S. G. Przhibel'skiĭ, and N. A. Chigir', Vestnik Moskov. Univ. Ser. Fiz. **33**(4), 35 (1978).

[7] S. G. Rautian and I. I. Sobelman, Zh. Éksp. Teor. Fiz. **41**, 456 (1961) [Sov. Phys. JETP **14**, 328 (1962)].

[8] T. Ya. Popova and A. K. Popov, Zh. Prikl. Spektrosk. **12**, 989 (1970); Izv. Vyssh. Uchebn. Zaved. Fiz. No. 11, 38 (1970).

[9] E. Y. Wu, S. Ezekiel, M. Ducloy *et al.*, Phys. Rev. Lett. **38**, 1077 (1977).

[10] I. M. Beterov and V. P. Chebotaev, JETP Lett. **9**, 127 (1969).

[11] Th. Hansch and P. Toschek, Z. Phys. **236**, 213 (1970).

[12] I. S. Zelikovich, S. A. Pul'kin, L. S. Gaĭda, and V. N. Komar, Zh. Éksp. Teor. Fiz. **94**, No. 12, 76 (1988) [Sov. Phys. JETP **67**, 2434 (1988)]; L. S. Gaĭda and S. A. Pul'kin, Opt. Spektrosk. **67**, 761 (1989) [Opt. Spectrosc. (USSR) **67**, 446 (1989)].

[13] T. Ya. Popova, A. K. Popov, S. G. Rautian, and R. I. Sokolovskiĭ, Zh. Éksp. Teor. Fiz. **57**, 850 (1969) [Sov. Phys. JETP **30**, 466 (1970)].

[14] *Atomic Coherence and Interference* (Crested Butte Workshop, 1993), Quantum Opt. **6, No. 4**, (1994); ICONO '95: *Coherent Phenomena and Amplification Without Inversion*, A. V. Andreev, O. Kocharovskaya, and P. Mandel (eds.), Proc. SPIE (1996), Vol. 2798; G. G. Padmabandu, G. R. Welch, I. N. Shubin *et al.*, Phys. Rev. Lett. **76**, 2053 (1996); G. Vemuri and G. S. Agarwal, Phys. Rev. A **53**, 1060 (1996).

[15] G. E. Notkin, S. G. Rautian, and A. A. Feoktistov, Zh. Éksp. Teor. Fiz. **52**, 1673 (1967) [Sov. Phys. JETP **25**, 1112 (1967)].

[16] T. Ya. Popova, A. K. Popov, S. G. Rautian, and A. A. Feoktistov, Zh. Éksp. Teor. Fiz. **57**, 444 (1969) [Sov. Phys. JETP **30**, 243 (1970)].

[17] M. S. Feld and A. Javan, Phys. Rev. **177**, 540 (1969); M. S. Feld, in *Fundamental and Applied Laser Physics* (Proc. of the Esfahan Symposium, August 29–September 5, 1971), M. S. Feld, A. Javan, and N. Kurnit, Wiley, New York (1973), p. 369.

[18] A. K. Popov and V. M. Shalaev, Opt. Commun. **35**, 189 (1980); Opt. Spektrosk. **49**, 617 (1980) [Opt. Spectrosc. (USSR) **49**, 336 (1980)].

[19] A. K. Popov, in *Proc. of Int. Conf. Lasers '79* (Orlando, Florida, 1979), STS Press, McLean, Va. (1980), p. 295; A. K. Popov and V. M. Shalaev, Appl. Phys. **27**, 63 (1981).

[20] A. K. Popov, Zh. Éksp. Teor. Fiz. **58**, 1623 (1970) [Sov. Phys. JETP **31**, 870 (1970)].

[21] De-Zhong Wang and Jin-Yue Gao, Phys. Rev. A **52**, 3201 (1995).

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Polarization fields in the positronium atom undergoing emission or absorption of optical photons

O. N. Gadomskiĭ and S. G. Moiseev

Ul'yanovsk State University, 432700 Ul'yanovsk, Russia

(Submitted 20 March 1997)
Zh. Éksp. Teor. Fiz. **113**, 471–488 (February 1998)

This paper solves the problem of the interaction of an electron and positron via the field of soft and hard photons with emission or absorption of a real photon. The interaction is interpreted as a third-order QED effect in the coordinate representation. The role of intermediate states with positive and negative frequencies is studied. A general expression is derived for the matrix elements of the operator of the effective electron–positron interaction energy for different types of quantum transitions. The expression makes it possible to calculate the probabilities of the corresponding transitions in the nonrelativistic approximation. Electric dipole transitions in the positronium atom accompanied by emission (absorption) of an optical photon are investigated. Two-particle wave functions of the positronium atom are used to introduce the concept of polarization fields inside the positronium atom. It is found that the polarization fields depend on the coordinates and time and on the choice of the pair of states between which a quantum transition with emission or absorption of a photon takes place. © *1998 American Institute of Physics.* [S1063-7761(98)00602-7]

1. INTRODUCTION

Polarization fields play an important role in shaping various optical processes. For instance, in the classical optics of insulators such a polarization field is the electric dipole field,[1] which allows not only for an explaination of reflection and refraction of light but also makes it possible to derive in a rigorous manner the Lorenz–Lorentz formula for the refractive index. As shown in Refs. 2 and 3, establishing the nature of polarization fields in insulators requires using an approach based on QED first- and second-order effects. This makes it possible to study various types of quantum transition in the field of virtual and real photons and to allow for the orbital and spin degrees of freedom of the atomic electrons, the various types of intermediate states in the spectrum of interacting atoms, and the retardation effect for atoms located at arbitrary distances from one another. A considerable achievement here is the development of a method for obtaining new types of integral equations that describe the propagation of photons in a medium with allowance for different quantum-transition types (electric dipole, quadrupole, magnetic dipole, spin, etc.) in the spectrum of atoms.[4,5] Here allowance is made for electron and positron polarization fields, which correspond to intermediate states of the atoms with positive and negative energies, respectively. An integral field equation in the electric dipole approximation has been used to study the laws of reflection and refraction of light in quantum[6] and nonlinear[4] optics, to build a theory of a nonlinear refractive index,[4] and to theoretically predict the near-field effect with allowance for a discrete distribution of the atoms near the the observation point.[4,7]. In contrast to Refs. 1–7, our paper studies the role of polarization fields inside an individual atom.

Drake[8] examined the interaction of two atomic electrons in a helium-like atom as a QED third-order effect. It was found that with allowance for the orbital and spin degrees of freedom and intermediate states with positive and negative energies, third-order effects lead only to relativistic corrections to the photon emission and absorption probability.

The present paper is devoted to the theory of quantum transitions between the levels of the positronium atom induced by soft (optical or radio-frequency) photons. The need for such a theory derives from the use of precision methods of radio and optical spectroscopy in the physics of the positronium atom, research into the annihilation process in highly excited states,[9–11] and the study of the possibility of markedly changing the kinetic characteristics of the annihilation process in the field of an optical laser.[12]. Obviously, the interaction of the positronium atom and the photon field is largely determined by the coupling constants in the effective Hamiltonian. Hence we pay special attention to these constants by interpreting the optical transition between the levels of the positronium atom as a QED third-order effect. We find that these effects explain the one-photon processes of emission and absorption in the positronium atom due to the Coulomb electron–positron interaction and induced polarization fields. We also find that calculations of the probability (per unit time) of spontaneous photon emission yield the same result if we use the single-particle wave functions of the positronium atom of Refs. 10 and 13 or the two-particle wave functions derived in the present paper. In other words, thanks to the specific properties of the positronium atom, third-order effects are most evident in such an atom, in contrast to the case of helium-like atoms.[8]

As noted earlier, in this paper we use the two-particle wave functions of the positronium atom. First, this makes it possible to study the various schemes of quantum transitions

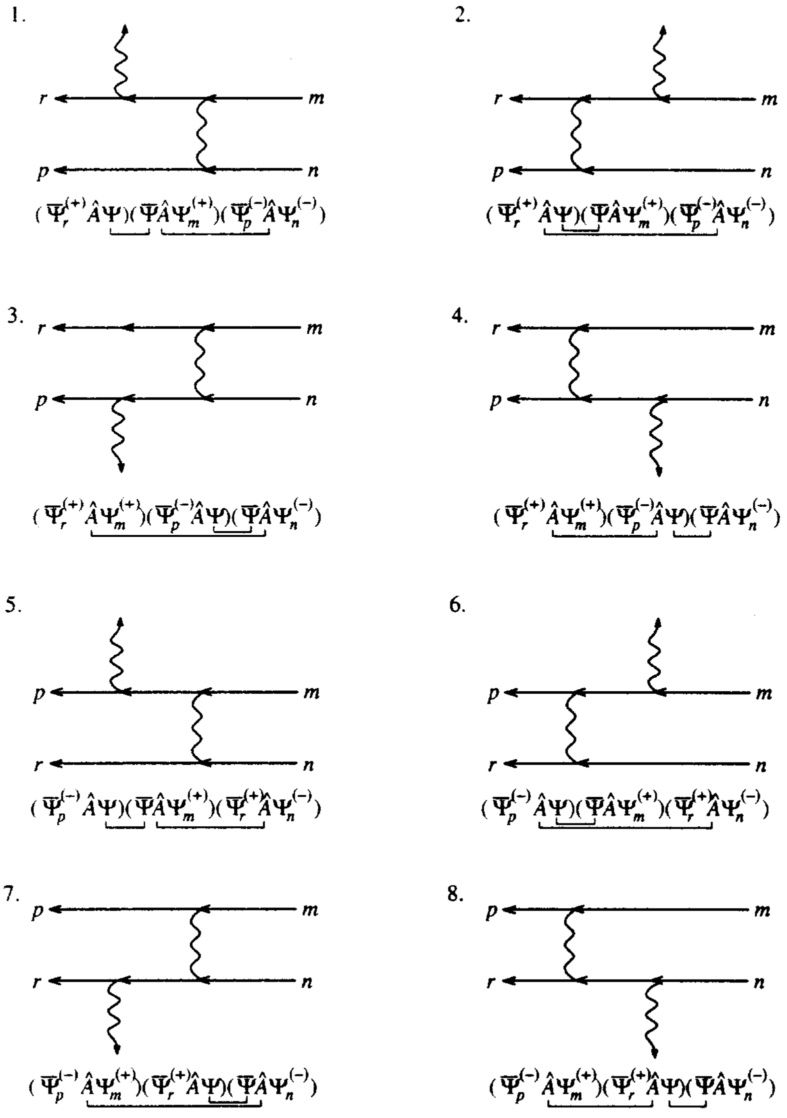


FIG. 1. Feynman diagrams for the electron–positron interaction in the positronium atom with emission or absorption of a photon.

accompanying the emission or absorption of an optical photon and to establish the role of real and virtual photons in such processes and the role of various intermediate states. Second, by employing two-particle wave functions we are able to describe the radiative interaction of the positronium atom in the field of annihilation and optical photons without resorting to perturbative techniques,[12] i.e., solely on the basis of fermion and boson operators.

Photon emission in electron–electron and electron–positron collisions was studied by Lifshitz,[14], Fedushin,[15] and Garibyan[16] under the stringent condition that the interacting particles have well-defined momenta before and after collision. This is true for the interaction of free particles, but is of no use if we wish to study quantum transitions between states of the positronium atom, where the constants of motion are the square of the total angular momentum (orbital and spin) and the projection of the total angular momentum on the quantization axis. Furthermore, as we show in this paper, intermediate states with positive and negative frequencies play an important role in the electron–positron interactions involving emission (absorption) of a photon. A study of the role of intermediate states provides us with certain capabilities for exciting the Rydberg states of the positronium atom by using, for example, the $2P$ state as an intermediate state.[9] In this paper, we therefore investigate the entirely new problem of the interaction of the electron and positron in a positronium atom undergoing emission (absorption) of a photon as a QED third-order effect. As noted above, the interaction of two atomic electrons belonging to two hydrogen-like atoms located at an arbitrary distance from each other was examined in Refs. 2 and 3 in the context of third-order effects. Here we use a similar approach to study the electron–positron interaction, but focus on important differences between our problem and the one discussed in Refs. 2 and 3, which involve, for example, the emergence of a hard electron–positron interaction mode due to the exchange of virtual gamma photons.

2. EFFECTIVE INTERACTION ENERGY OF A POSITRONIUM ATOM AND THE PHOTON FIELD

In quantum transition theory one can introduce the effective Hamiltonian of interacting electron–positron and electromagnetic fields, which contains the term

$$G(\mathbf{k})a_2^\dagger b_2^\dagger b_1 a_1 c + G^*(-\mathbf{k})a_1^\dagger b_1^\dagger b_2 a_2 c^\dagger, \tag{2.1}$$

where the subscripts 1 and 2 label the quantum states of the positronium atom, $b_1 a_1$ is the annihilation operator for the positronium atom in state 1, $a_2^\dagger b_2^\dagger$ is the creation operator for the positronium atom in state 2, etc., and c and $c^\dagger$ are the annihilation and creation operators for a photon with wave vector $\mathbf{k}$. The interaction (2.1) is considered a QED third-order effect, with corresponding Feynman diagrams depicted in Fig. 1.

Separating the temporal factors from the wave functions and integrating in the S-matrix $S_{i\to f}^{(3)}$ with respect to time, frequencies, and wave vectors, we arrive at the following expression for the matrix of the effective interaction energy ($\hbar = c = 1$):

$$\begin{aligned}
U_{i\to f}^{(3)} = {} & \frac{e^3}{4\pi} a_2^\dagger b_2^\dagger b_1 a_1 \int d\mathbf{r}' d\mathbf{r}'' d\mathbf{r}''' \\
& \times \Bigg\{ -\frac{1}{|\mathbf{r}''-\mathbf{r}'''|} \sum_{l_\pm} \frac{\exp(i|\omega_2^{(-)}-\omega_1^{(-)}||\mathbf{r}''-\mathbf{r}'''|)}{\omega_l(1-i0)+\omega-\omega_2^{(+)}} \\
& \times \bar{\Psi}_2^{(+)}(\mathbf{r}')\hat{A}(\mathbf{r}')\Psi_l(\mathbf{r}')\bar{\Psi}_l(\mathbf{r}'')\gamma_\mu'' \\
& \times \Psi_1^{(+)}(\mathbf{r}'')\bar{\Psi}_1^{(-)}(\mathbf{r}''')\gamma_\mu''\Psi_2^{(-)}(\mathbf{r}''') \\
& -\frac{1}{|\mathbf{r}'-\mathbf{r}'''|} \sum_{l_\pm} \frac{\exp(i|\omega_1^{(-)}-\omega_2^{(-)}||\mathbf{r}'-\mathbf{r}'''|)}{\omega_l(1-i0)-\omega-\omega_1^{(+)}} \\
& \times \bar{\Psi}_2^{(+)}(\mathbf{r}')\gamma_\mu'\Psi_l(\mathbf{r}')\bar{\Psi}_l(\mathbf{r}'')\hat{A}(\mathbf{r}'') \\
& \times \Psi_1^{(+)}(\mathbf{r}'')\bar{\Psi}_1^{(-)}(\mathbf{r}''')\gamma_\mu'''\Psi_2^{(-)}(\mathbf{r}''') \\
& -\frac{1}{|\mathbf{r}'-\mathbf{r}'''|} \sum_{l_\pm} \frac{\exp(i|\omega_2^{(+)}-\omega_1^{(+)}||\mathbf{r}'-\mathbf{r}'''|)}{\omega_l(1-i0)+\omega-\omega_1^{(-)}} \\
& \times \bar{\Psi}_2^{(+)}(\mathbf{r}')\gamma_\mu'\Psi_1^{(+)}(\mathbf{r}')\bar{\Psi}_1^{(-)}(\mathbf{r}'')\hat{A}(\mathbf{r}'') \\
& \times \Psi_l(\mathbf{r}'')\bar{\Psi}_l(\mathbf{r}''')\gamma_\mu'''\Psi_2^{(-)}(\mathbf{r}''') \\
& -\frac{1}{|\mathbf{r}'-\mathbf{r}''|} \sum_{l_\pm} \frac{\exp(i|\omega_2^{(+)}-\omega_1^{(+)}||\mathbf{r}'-\mathbf{r}''|)}{\omega_l(1-i0)-\omega-\omega_2^{(-)}} \\
& \times \bar{\Psi}_2^{(+)}(\mathbf{r}')\gamma_\mu'\Psi_1^{(+)}(\mathbf{r}')\bar{\Psi}_1^{(-)}(\mathbf{r}'')\gamma_\mu'' \\
& \times \Psi_l(\mathbf{r}'')\bar{\Psi}_l(\mathbf{r}''')\hat{A}(\mathbf{r}''')\Psi_2^{(-)}(\mathbf{r}''') \\
& +\frac{1}{|\mathbf{r}''-\mathbf{r}'''|} \sum_{l_\pm} \frac{\exp(i|\omega_2^{(-)}-\omega_1^{(+)}||\mathbf{r}''-\mathbf{r}'''|)}{\omega_l(1-i0)+\omega-\omega_2^{(-)}} \\
& \times \bar{\Psi}_2^{(-)}(\mathbf{r}')\hat{A}(\mathbf{r}')\Psi_l(\mathbf{r}')\bar{\Psi}_l(\mathbf{r}'')\gamma_\mu'' \\
& \times \Psi_1^{(+)}(\mathbf{r}'')\bar{\Psi}_1^{(+)}(\mathbf{r}''')\gamma_\mu'''\Psi_2^{(-)}(\mathbf{r}''') \\
& +\frac{1}{|\mathbf{r}'-\mathbf{r}'''|} \sum_{l_\pm} \frac{\exp(i|\omega_2^{(-)}-\omega_1^{(+)}||\mathbf{r}'-\mathbf{r}'''|)}{\omega_l(1-i0)-\omega-\omega_1^{(+)}} \\
& \times \bar{\Psi}_2^{(-)}(\mathbf{r}')\gamma_\mu'\Psi_l(\mathbf{r}')\bar{\Psi}_l(\mathbf{r}'')\hat{A}(\mathbf{r}'') \\
& \Psi_1^{(+)}(\mathbf{r}'')\bar{\Psi}_1^{(+)}(\mathbf{r}''')\gamma_\mu'''\Psi_2^{(-)}(\mathbf{r}''') \\
& +\frac{1}{|\mathbf{r}'-\mathbf{r}'''|} \sum_{l_\pm} \frac{\exp(i|\omega_2^{(-)}-\omega_1^{(+)}||\mathbf{r}'-\mathbf{r}'''|)}{\omega_l(1-i0)+\omega-\omega_1^{(+)}} \\
& \times \bar{\Psi}_2^{(-)}(\mathbf{r}')\gamma_\mu'\Psi_1^{(+)}(\mathbf{r}')\bar{\Psi}_1^{(+)}(\mathbf{r}'')\hat{A}(\mathbf{r}'') \\
& \times \Psi_l^{(+)}(\mathbf{r}'')\bar{\Psi}_l(\mathbf{r}''')\gamma_\mu'''\Psi_2^{(-)}(\mathbf{r}''') \\
& +\frac{1}{|\mathbf{r}'-\mathbf{r}''|} \sum_{l_\pm} \frac{\exp(i|\omega_2^{(-)}-\omega_1^{(+)}||\mathbf{r}'-\mathbf{r}''|)}{\omega_l(1-i0)-\omega-\omega_2^{(-)}} \\
& \times \bar{\Psi}_2^{(-)}(\mathbf{r}')\gamma_\mu'\Psi_1^{(+)}(\mathbf{r}')\bar{\Psi}_1^{(+)}(\mathbf{r}'')\gamma_\mu'' \\
& \times \Psi_l^{(+)}(\mathbf{r}'')\bar{\Psi}_l(\mathbf{r}''')\hat{A}(\mathbf{r}''')\Psi_2^{(-)}(\mathbf{r}''') \Bigg\},
\end{aligned} \tag{2.2}$$

where $\Psi_{1(2)}^{(\pm)}(\mathbf{r})$ are the solutions of the Dirac equation for an electron with positive and negative frequencies in the positronium atom, which may be in quantum states 1 and 2; $\bar{\Psi}_{1(2)}^{(\pm)} = \Psi_{1(2)}^{(\pm)*}\gamma_4$, $\Psi_{1(2)}^{(\pm)*}$ is the conjugate wave function, $\gamma_4 = \beta$, $\gamma_j = -i\beta\alpha_j$ $(j = 1,2,3)$,

$$\alpha = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and $\boldsymbol{\sigma}$ are the Pauli matrices. The primes on the radius vectors $\mathbf{r}$ and the γ matrices correspond to different wave functions of the interacting particles. Here the γ matrices with different numbers of primes commute. Summation in (2.2) is over all intermediate states with positive and negative frequencies. Electron states with negative frequencies are interpreted as positron states, and we do not introduce the positron wave function, which contains the charge conjugation transformation. The reason is that a linear combination of states with opposite signs of charge cannot be a general solution of the Dirac equation.[13] In this paper we use the solution

$$\begin{aligned}
\Psi &= \sum_r a_r \Psi_r^{(+)} + \sum_r b_r^\dagger \Psi_r^{(-)}, \\
\bar{\Psi} &= \sum_r a_r^\dagger \bar{\Psi}_r^{(+)} + \sum_r b_r \bar{\Psi}_r^{(-)}.
\end{aligned} \tag{2.2a}$$

We transform from $S_{i\to j}^{(3)}$ to the matrix of the effective interaction energy in (2.2) according to the following equality:

$$S_{i\to j}^{(3)} = -2\pi i U_{i\to j}^{(3)} \delta(-\omega_m^{(+)} - \omega + \omega_r^{(+)} + \omega_n^{(-)} - \omega_p^{(-)}), \tag{2.3}$$

where $\omega_m^{(+)}$, $\omega_r^{(+)}$, $\omega_n^{(-)}$, and $\omega_p^{(-)}$ are the frequencies of the interacting electron and positron in the positronium atom. Here, by fixing the states of the positronium atom 1 and 2, we have $r = 2$, $p = 1$, $m = 1$, and $n = 2$ for diagrams 1–4 and the corresponding terms in (2.2). For diagrams 5–8 we have $r = 1$, $p = 2$, $m = 1$, and $n = 2$. The sign of the frequency ω of the optical photon in the conservation law (2.3) indicates that in the vector potential operator $\hat{A} = \Sigma\gamma_\mu A_\mu$ we have specified the negative-frequency part, which is proportional to the annihilation operator for a photon of this mode.

Integration in the matrix $S_{i\to j}^{(3)}$ with respect to time, frequencies, and wave vectors makes it possible to find the frequencies of the virtual photons, which are carriers of the electron–positron interaction. Here, diagrams 1–4 in Fig. 1 correspond to electron–positron interaction via soft virtual photons, whose energy is much less than twice the electron rest mass. We call this mode of electron–positron interaction *soft*. Diagrams 5–8 correspond to the hard mode of electron–positron interaction, with the virtual photons having an energy comparable to twice the electron rest mass. Moreover, the presence in (2.2) of factors containing frequency differences in their denominators suggests that the electron–positron interaction is resonant, in which case diagrams 1–4 of the soft interaction mode and the diagrams 5–8 of the hard interaction mode correspond to processes that involve intermediate states with both positive (l_+) and negative (l_-) frequencies. The terms that provide the greatest contribution to (2.2) are those whose denominators contain frequency differences much less than twice the electron rest mass.

2.1. Generalized Breit operator

The Breit operator corresponds to QED second-order effects,[13] and was derived for the interaction of two free electrons and for the atomic electrons in helium-like atoms. Pirenne[17] and Berestetskiĭ and Landau[18] generalized this operator to the electron–positron interaction with allowance for the charge conjugation transformation and the exchange interaction between the electron and positron. In Refs. 2, 3, and 19, the Breit operator was also generalized to the case of two atomic electrons belonging to two hydrogen-like atoms an arbitrary distance from each other. The electron–positron interaction considered in the present paper contains the exchange of virtual photons as an integral part of a more complicated interaction, with Feynman diagrams depicted in Fig. 1. We now show how allowing for this exchange in QED third-order effects for the electron–positron interaction leads to a generalization of the Breit operator, where in addition to retardation effects we must properly allow for the role of intermediate states and the nonrelativistic motion of the interacting particles in the positronium atom.

Let us examine the first term in the matrix (2.2) of the effective interaction energy and expand the retardation factor in this term to order $1/c^2$. Assuming that $c\neq 1$, we obtain

$$\frac{\exp((i/c)|\omega_2^{(-)}-\omega_1^{(-)}||\mathbf{r}''-\mathbf{r}'''|)}{|\mathbf{r}''-\mathbf{r}'''|} = \frac{1}{|\mathbf{r}''-\mathbf{r}'''|}+i\frac{|\omega_2^{(-)}-\omega_1^{(-)}|}{c}-\frac{(\omega_2^{(-)}-\omega_1^{(-)})^2}{2c^2}|\mathbf{r}''-\mathbf{r}'''|. \tag{2.4}$$

Substituting (2.4) into the first term in (2.2), we find that the second term on the right-hand side of Eq. (2.4) contributes nothing to the first term in (2.2), since the wave functions $\Psi_l,\Psi_1^{(+)}$ and $\Psi_1^{(-)},\Psi_2^{(-)}$ are orthogonal. Furthermore, since $\gamma_4''\gamma_\mu''\gamma_4'''\gamma_\mu'''=1-\alpha''\alpha'''$ and the matrix elements of the operators α'' and α''' are equal in order of magnitude to v/c (v is the particle velocity), in terms containing $\alpha''\alpha'''$ we can keep just the first term on the right-hand side of Eq. (2.4).

We remove the frequencies from the expression for the matrix element by employing the Dirac equation:

$$\begin{aligned} &H''\Psi_1^{(+)}(\mathbf{r}'')=\omega_1^{(+)}\Psi_1^{(+)}(\mathbf{r}''),\\ &H''=c\alpha''\mathbf{p}''+\gamma_4''mc^2-\frac{e^2}{2|\mathbf{r}''-\mathbf{r}'''|},\\ &H'''\Psi_2^{(-)}(\mathbf{r}''')=\omega_2^{(-)}\Psi_2^{(-)}(\mathbf{r}'''),\\ &H'''=c\alpha'''\mathbf{p}'''+\gamma_4'''mc^2-\frac{e^2}{2|\mathbf{r}''-\mathbf{r}'''|}, \end{aligned} \tag{2.5}$$

where m is the electron mass, and $\mathbf{p}''$ and $\mathbf{p}'''$ are the momentum operators acting on the position vectors $\mathbf{r}''$ and $\mathbf{r}'''$, respectively. Now we write the obvious relationship

$$-(\omega_2^{(-)}-\omega_1^{(-)})^2|\mathbf{r}''-\mathbf{r}'''| = \frac{\omega_2^{(-)}-\omega_1^{(-)}}{\omega_l-\omega_1^{(+)}}[H'',[H''',|\mathbf{r}''-\mathbf{r}'''|]]. \tag{2.6}$$

Calculating the commutators, we find that the contribution of the third term on the right-hand side of Eq. (2.4) is determined by the operator

$$-\frac{(\omega_2^{(-)}-\omega_1^{(-)})^2}{2c^2}|\mathbf{r}''-\mathbf{r}'''|\to\frac{(\alpha''\alpha''')-(\alpha''\mathbf{n}_{23})(\alpha'''\mathbf{n}_{23})}{2|\mathbf{r}''-\mathbf{r}'''|}\times\frac{\omega_2^{(-)}-\omega_1^{(-)}}{\omega_l-\omega_1^{(+)}}, \tag{2.7}$$

where $\mathbf{n}_{23}=(\mathbf{r}''-\mathbf{r}''')/|\mathbf{r}''-\mathbf{r}'''|$. Thus, the electron–positron interaction corresponding to diagram 1 in Fig. 1 is given by the generalized Breit operator

$$B_{1l}(\mathbf{r}'',\mathbf{r}''')=\frac{e^2}{4\pi}\frac{1}{|\mathbf{r}''-\mathbf{r}'''|}+\frac{e^2}{4\pi}\frac{(\alpha''\alpha''')-(\alpha''\mathbf{n}_{23})(\alpha'''\mathbf{n}_{23})}{2|\mathbf{r}''-\mathbf{r}'''|}f_{l1}-\frac{e^2}{4\pi}\frac{(\alpha''\alpha''')}{|\mathbf{r}''-\mathbf{r}'''|}, \tag{2.8}$$

where $f_{l1}=(\omega_2^{(-)}-\omega_1^{(-)})/(\omega_{l_+}-\omega_1^{(+)})$. At $f_{l1}=1$ the operator (2.8) is the relativistic Breit operator.[13] The condition $f_{l1}=1$ means that the transition of the system of two particles into an intermediate state must obey particle energy conservation, i.e., $\omega_2^{(-)}-\omega_1^{(-)}=\omega_l-\omega_1^{(+)}$.

Similarly, we can examine the exchange of virtual photons in the second, third, and fourth terms of the matrix (2.2), with diagrams 2, 3, and 4, respectively. Performing the necessary calculations, we obtain the operators $B_{2l}(\mathbf{r}',\mathbf{r}''')$, $B_{3l}(\mathbf{r}',\mathbf{r}''')$, and $B_{4l}(\mathbf{r}',\mathbf{r}'')$, which are similar to (2.8) but with f_{l1} replaced by the following coefficients:

$$f_{l2}=\frac{\omega_2^{(-)}-\omega_1^{(-)}}{\omega_2^{(+)}-\omega_{l_+}},\qquad f_{l3}=\frac{\omega_1^{(+)}-\omega_2^{(+)}}{\omega_{l_-}-\omega_2^{(-)}},$$

$$f_{l4}=\frac{\omega_1^{(+)}-\omega_2^{(+)}}{\omega_1^{(-)}-\omega_{l_-}}. \tag{2.9}$$

Thus, the contribution of the various diagrams 1–4 to the electron–positron interaction depends heavily on the type of positive- or negative-frequency intermediate state.

We now examine diagrams 5–8, which differ from 1–4 in that r is replaced by p and p by r. The retardation factor in the corresponding terms of the matrix (2.2) has a large coefficient $2mc$ in the exponential and hence oscillates with a wavelength $1/2mc$ ($1/2mc$ is the electron Compton wavelength). Allowing for the important role of intermediate states in the electron–positron interaction, noted earlier in connection with diagrams 1–4, we employ the fact that the most important domain of integration for diagrams 5–8 is $|\mathbf{r}''-\mathbf{r}'''|\sim 1/mc$. Since according to our assumption that the electron and positron velocities in the positronium atom are small, the electron and positron wave functions change little in this region, in (2.2) we can separate out integration with respect to $\mathbf{r}=\mathbf{r}''-\mathbf{r}'''$, putting $\mathbf{r}''=\mathbf{r}'''$ in the arguments of the wave functions. Then the fifth term in (2.2) takes the form

$$\begin{aligned}&\frac{e^3}{4\pi}a_2^\dagger b_2^\dagger a_1 b_1\int d\mathbf{r}'\,d\mathbf{r}''\,d\mathbf{r}'''\,\frac{\exp(2imcr)}{2}\\&\quad\times\sum_{l_\pm}\frac{1}{\omega_l(1-i0)+\omega-\omega_2^{(-)}}\,\overline{\Psi}_2^{(-)}(\mathbf{r}')\hat{A}(\mathbf{r}')\Psi_l(\mathbf{r}')\\&\quad\times\{\Psi_l^*(\mathbf{r}'')\Psi_1^{(+)*}(\mathbf{r}'')\Psi_1^{(+)}(\mathbf{r}'')\Psi_2^{(-)}(\mathbf{r}'')\\&\quad-(\Psi_l^*(\mathbf{r}'')\alpha''\Psi_1^{(+)}(\mathbf{r}''))(\Psi_1^{(+)*}(\mathbf{r}'')\alpha'''\Psi_2^{(-)}(\mathbf{r}''))\}.\end{aligned} \tag{2.10}$$

The improper integral over $\mathbf{r}$ is equal to $-\pi/(mc)^2$. Thus, the electron–positron interaction corresponding to the hard interaction mode of Feynman diagram 5 is determined by the operator[13]

$$B_{5l}(\mathbf{r}'',\mathbf{r}''')=-\frac{e^2}{4\pi}\frac{\pi}{m^2c^2}(1-\alpha''\alpha''')\,\delta(\mathbf{r}''-\mathbf{r}'''). \tag{2.11}$$

The same operator can be separated out of the other terms of the matrix (2.2) that correspond to the hard mode of the electron–positron interaction. The dependence of the operator (2.11) on the dummy indices is only meaningful when (2.11) is substituted into the appropriate matrix element of (2.2).

2.2. Allowance for other intermediate states

In the generalized Breit operators B_{sl} ($s=1,2,3,4$) and in the operators B_{5l}, we allowed only for intermediate states that correspond to the maximum values of the resonant factors, which contain differences of various frequencies. We now take into account the remaining intermediate states, for which the resonant factors become equal to $\pm 1/2mc^2$. Then for the soft mode of the electron–positron interaction, we can associate with the effective interaction energy matrix an effective interaction energy operator that allows for emission (absorption) of a real photon. This operator has the form

$$\begin{aligned}U_{\text{soft}}^{(3)}=-\frac{e^3}{8\pi mc^2}\biggl\{&-\gamma_4'\gamma_\delta' A_\delta(\mathbf{r}')\frac{1}{|\mathbf{r}''-\mathbf{r}'''|}\\&+\gamma_4'''\gamma_\delta''' A_\delta(\mathbf{r}''')\frac{1}{|\mathbf{r}'-\mathbf{r}''|}\biggr\},\end{aligned} \tag{2.12}$$

where we have left only the Coulomb term in the generalized Breit operators, since the coefficients f_{l1}, f_{l2}, f_{l3}, and f_{l4} contain a factor $1/c^2$.

3. TRANSITION TO TWO-COMPONENT ELECTRON AND POSITRON WAVE FUNCTIONS

When the particle velocities are low ($v\ll c$), we can go from four-component wave functions $\Psi_n^{(\pm)}$ to approximate two-component wave functions $\Phi_n^{(\pm)}$.[13] This transition is accurate to order $1/c^2$. We write the functions $\Psi_n^{(\pm)}$ in the following way:

$$\begin{aligned}&\Psi_n^{(+)}=\begin{pmatrix}\varphi_n^{(+)}\\ \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{2mc}\varphi_n^{(+)}\end{pmatrix},\qquad \varphi_n^{(+)}=\left(1-\frac{p^2}{8m^2c^2}\right)\Phi_n^{(+)},\\&\Psi_n^{(-)}=\begin{pmatrix}-\dfrac{\sigma\cdot\mathbf{p}}{2mc}\chi_n^{(-)}\\ \chi_n^{(-)}\end{pmatrix},\qquad \chi_n^{(-)}=\left(1-\frac{p^2}{8m^2c^2}\right)\Phi_n^{(-)},\end{aligned} \tag{3.1}$$

where the operators $\boldsymbol{\sigma}$ and $\mathbf{p}$ act on the variables on which the wave functions $\Phi_n^{(\pm)}$ depend. Henceforth we use labels 1, 2, and 3 for the corresponding operators instead of primes.

Using (3.1), we transform the matrix element $\langle\overline{\Psi}_2^{(+)}\times(\mathbf{r}')\hat{A}(\mathbf{r}')\Psi_l^{(+)}(\mathbf{r}')\rangle$ in the first term on the right-hand side of Eq. (2.2), whereupon we obtain, to order $1/c^3$,

$$\begin{aligned}&\langle\overline{\Psi}_2^{(+)}\hat{A}\Psi_l^{(+)}\rangle=\langle\Phi_2^{(+)}(\mathbf{r}')|U_{NR}^{(1)}|\Phi_l^{(+)}(\mathbf{r}')\rangle,\\&U_{NR}^{(1)}=\frac{1}{mc}(\mathbf{p}_1\cdot\mathbf{A}_1)+\frac{\hbar}{2mc}(\boldsymbol{\sigma}_1\cdot\mathbf{H}_1)\\&\quad+\frac{1}{16m^3c^3}\{2(\mathbf{A}_1\cdot\mathbf{p}_1)\mathbf{p}_1^2+2\mathbf{p}_1^2(\mathbf{A}_1\cdot\mathbf{p}_1)\\&\quad+\hbar(\boldsymbol{\sigma}_1\cdot\mathbf{H}_1)\mathbf{p}_1^2+\hbar\mathbf{p}_1^2(\boldsymbol{\sigma}_1\cdot\mathbf{H}_1)\},\end{aligned} \tag{3.2}$$

where we have allowed for the fact that $\nabla_1\cdot\mathbf{A}_1=0$, and $\mathbf{H}_1=\nabla_1\times\mathbf{A}_1$ is the magnetic field strength at the point $\mathbf{r}'$. Hereafter, we use the conventional Gaussian system of units and introduce the constant $\hbar$.

Similarly, using (3.1) we can transform $\langle\Psi_2^{(-)}\hat{A}\Psi_l^{(-)}\rangle$. After performing the necessary transition, we find the same nonrelativistic operator as in (3.2), which transforms the wave function $\Phi_l^{(-)}$ into the function $\Phi_2^{(-)}$.

We now transform $\langle \Psi_{l_+}^{*}(\mathbf{r}'')\Psi_1^{(-)*}(\mathbf{r}''')|B_{1l}|\Psi_1^{(+)}\times(\mathbf{r}'')\Psi_2^{(-)}(\mathbf{r}''')\rangle$, the matrix element in the first term on the right-hand side of Eq. (2.2), to order $1/c^2$. Here the Coulomb term in the operator (2.8) is transformed using the wave functions $\varphi_n^{(\pm)}$ and $\chi_n^{(\pm)}$, which contain the operator $1-p^2/8m^2c^2$. To transform the retarded terms in (2.8), we simply replace $\varphi_n^{(+)}$ and $\chi_n^{(-)}$ by $\Phi_n^{(+)}$ and $\Phi_n^{(-)}$.

Omitting the uninteresting computational details, we arrive at the following nonrelativistic operator, which transforms the wave function $\Phi_1^{(+)}(\mathbf{r}'')\Phi_2^{(-)}(\mathbf{r}''')$ of a pair of particles into the wave function $\Phi_l^{(+)}(\mathbf{r}'')\Phi_1^{(-)}(\mathbf{r}''')$:

$$
\begin{aligned}
B_{1l}(\mathbf{r}'',\mathbf{r}''') = {} & \frac{e^2}{4\pi}\,\frac{1}{|\mathbf{r}''-\mathbf{r}'''|} \\
& -\frac{e^2}{4\pi}\,\frac{1}{4m^2c^2}\left(\frac{1}{2}f_{fl}-1\right)\frac{4}{|\mathbf{r}''-\mathbf{r}'''|}(\mathbf{p}_2\cdot\mathbf{p}_3) \\
& +\frac{1}{2}f_{l1}\,\frac{1}{4m^2c^2}\,\frac{e^2}{4\pi}\,\frac{4}{|\mathbf{r}''-\mathbf{r}'''|}\mathbf{n}_{23}\cdot(\mathbf{n}_{23}\cdot\mathbf{p}_3)\mathbf{p}_2 \\
& +\frac{e^2}{4\pi}\,\frac{\hbar}{4m^2c^2}\,\frac{1}{|\mathbf{r}''-\mathbf{r}'''|^2} \\
& \times\Bigg\{\boldsymbol{\sigma}_3\cdot(\mathbf{n}_{23}\times\mathbf{p}_3)-\boldsymbol{\sigma}_2\cdot(\mathbf{n}_{23}\times\mathbf{p}_2) \\
& +4\boldsymbol{\sigma}_3\cdot(\mathbf{n}_{23}\times\mathbf{p}_1)\left(\frac{1}{2}f_{l1}-1\right) \\
& -2\boldsymbol{\sigma}_2\cdot(\mathbf{n}_{23}\times\mathbf{p}_3)\left(\frac{1}{2}f_{l1}-1\right) \\
& +f_{l1}\boldsymbol{\sigma}_2\cdot(\mathbf{n}_{23}\times\mathbf{p}_3)\Bigg\} \\
& +\frac{e^2}{4\pi}\,\frac{1}{2m^2c^2}\,\frac{\hbar^2}{|\mathbf{r}''-\mathbf{r}'''|^3}\left(\frac{1}{2}f_{l1}-1\right) \\
& \times\{(\boldsymbol{\sigma}_2\cdot\boldsymbol{\sigma}_3)-3(\boldsymbol{\sigma}_2\cdot\mathbf{n}_{23})(\boldsymbol{\sigma}_3\cdot\mathbf{n}_{23})\} \\
& +\pi\frac{e^2}{4\pi}\,\frac{\hbar^2}{m^2c^2}\,\delta(\mathbf{r}''-\mathbf{r}''') \\
& -\frac{e^2}{4\pi}\,\frac{\hbar^2}{2m^2c^2}\left(\frac{1}{2}f_{l1}-1\right) \\
& \times\frac{8\pi}{3}(\boldsymbol{\sigma}_2\cdot\boldsymbol{\sigma}_3)\,\delta(\mathbf{r}''-\mathbf{r}'''),
\end{aligned}
\tag{3.3}
$$

where $\mathbf{n}_{23}=(\mathbf{r}''-\mathbf{r}''')/|\mathbf{r}''-\mathbf{r}'''|$. Since the operator (3.3) contains higher orders of $1/|\mathbf{r}''-\mathbf{r}'''|$, we must remove the singularity by isolating the origin $\mathbf{r}''=\mathbf{r}'''$. The integral over the surface surrounding the origin remains finite as $\mathbf{r}''\to\mathbf{r}'''$. At $f_{l1}=1$, the operator (3.3) becomes the Breit operator for an electron and positron.

The transformations of the second, third, and fourth terms in (2.2) for the soft electron–positron interaction mode yield the same nonrelativistic operator as in (3.3), with the labels on the operators of momentum, spin, and radius vector replaced appropriately.

3.1. Hard mode of the electron–positron interaction

Employing the wave functions (3.1), we transform $\langle\overline{\Psi}_l^{(-)}(\mathbf{r}'')\overline{\Psi}_1^{(+)}(\mathbf{r}''')|B_{5l}|\Psi_1^{(+)}(\mathbf{r}'')\Psi_2^{(-)}(\mathbf{r}''')\rangle$, the matrix element in the fifth term on the right-hand side of Eq. (2.2).

Allowing for the fact that the operator (2.11) already contains the factor $1/c^2$, we can drop the operator $O=1-p^2/8m^2c^2$ in the wave functions (3.1). Multiplying the wave functions in the desired matrix element, we find that the matrix element of the first term in the operator (2.11) is a quantity of order $1/c^4$, i.e., it lies beyond the approximation scheme adopted here. The matrix element of the second term in (2.11) yields the operator

$$
B_{5l}(\mathbf{r}'',\mathbf{r}''')=\frac{\pi}{2}\,\frac{e^2}{4\pi}\,\frac{1}{m^2c^2}(3+\boldsymbol{\sigma}_2\cdot\boldsymbol{\sigma}_3)\,\delta(\mathbf{r}''-\mathbf{r}'''), \tag{3.4}
$$

which coincides with the exchange interaction operator for an electron and positron.[13]

The expression (2.10) also contains a sum over the intermediate states l_+ of electrons with positive energies. This means that we must calculate matrix elements of the type $\langle\overline{\Psi}_l^{(+)}(\mathbf{r}'')\overline{\Psi}_1^{(+)}(\mathbf{r}''')|B_{5l}|\Psi_1^{(+)}(\mathbf{r}'')\Psi_2^{(-)}(\mathbf{r}''')\rangle$. Here, however, the denominator of the resonant factor is a quantity of order $2mc^2$, so that matrix elements of this type can be discarded.

We now examine the sixth term in (2.2). To this end, we use the wave functions (3.1) to transform the matrix element $\langle\overline{\Psi}_2^{(-)}(\mathbf{r}')\overline{\Psi}_1^{(+)}(\mathbf{r}''')|B_{6l}|\Psi_l^{(+)}(\mathbf{r}')\Psi_2^{(-)}(\mathbf{r}''')\rangle$. We can easily show that the operator B_{6l} coincides with (3.4) if the labels are replaced accordingly.

3.2. Operator for the effective interaction energy of an electron and positron with a field of virtual and real photons

We substitute the operators (3.4), (3.3), and $U_{NR}^{(1)}$ into the matrix (2.2). Then in (2.2) we separate out terms with intermediate states ω_l containing denominators in the resonant factors that are much smaller than $2m^2c^2$. We denote this part of (2.2) by $U_{i\to f}^{(3)}\left(\begin{smallmatrix} l_+ \\ l_- \end{smallmatrix}\right)$. We then find that

$$
\begin{aligned}
U^{(+)}_{i\to f}\begin{pmatrix} l_+ \\ l_- \end{pmatrix} = \frac{1}{\hbar} a_2^\dagger b_2^\dagger b_1 a_1 \Bigg\{ & -\sum_{l_+} \frac{\langle \Phi_2^{(+)} | U_{NR}^{(1)}(\mathbf{r}') | \Phi_l^{(+)} \rangle \langle \Phi_l^{(+)} \Phi_1^{(-)} | B_{1l}(\mathbf{r}'',\mathbf{r}''') | \Phi_1^{(+)} \Phi_2^{(-)} \rangle}{\omega_l(1-i0)+\omega-\omega_2^{(+)}} \\
& -\sum_{l_+} \frac{\langle \Phi_2^{(+)} \Phi_1^{(-)} | B_{2l}(\mathbf{r}',\mathbf{r}''') | \Phi_l^{(+)} \Phi_2^{(-)} \rangle \langle \Phi_l^{(+)} | U_{NR}^{(1)}(\mathbf{r}'') | \Phi_1^{(+)} \rangle}{\omega_l(1-i0)-\omega-\omega_1^{(+)}} \\
& -\sum_{l_-} \frac{\langle \Phi_1^{(-)} | U_{NR}^{(1)}(\mathbf{r}'') | \Phi_l^{(-)} \rangle \langle \Phi_2^{(+)} \Phi_l^{(-)} | B_{3l}(\mathbf{r}',\mathbf{r}''') | \Phi_1^{(+)} \Phi_2^{(-)} \rangle}{\omega_l(1-i0)+\omega-\omega_1^{(-)}} \\
& -\sum_{l_-} \frac{\langle \Phi_2^{(+)} \Phi_1^{(-)} | B_{4l}(\mathbf{r}',\mathbf{r}'') | \Phi_1^{(+)} \Phi_l^{(-)} \rangle \langle \Phi_l^{(-)} | U_{NR}^{(1)}(\mathbf{r}''') | \Phi_2^{(-)} \rangle}{\omega_l(1-i0)-\omega-\omega_2^{(-)}} \\
& +\sum_{l_-} \frac{\langle \Phi_2^{(-)} | U_{NR}^{(1)}(\mathbf{r}') | \Phi_l^{(-)} \rangle \langle \Phi_l^{(-)} \Phi_1^{(+)} | B_{5l}(\mathbf{r}'',\mathbf{r}''') | \Phi_1^{(+)} \Phi_2^{(-)} \rangle}{\omega_l(1-i0)+\omega-\omega_2^{(-)}} \\
& +\sum_{l_+} \frac{\langle \Phi_2^{(-)} \Phi_1^{(+)} | B_{6l}(\mathbf{r}',\mathbf{r}''') | \Phi_l^{(+)} \Phi_2^{(-)} \rangle \langle \Phi_l^{(+)} | U_{NR}^{(1)}(\mathbf{r}'') | \Phi_1^{(+)} \rangle}{\omega_l(1-i0)-\omega-\omega_1^{(+)}} \\
& +\sum_{l_+} \frac{\langle \Phi_1^{(+)} | U_{NR}^{(1)}(\mathbf{r}'') | \Phi_l^{(+)} \rangle \langle \Phi_2^{(-)} \Phi_l^{(+)} | B_{7l}(\mathbf{r}',\mathbf{r}''') | \Phi_1^{(+)} \Phi_2^{(-)} \rangle}{\omega_l(1-i0)+\omega-\omega_1^{(+)}} \\
& +\sum_{l_-} \frac{\langle \Phi_2^{(-)} \Phi_1^{(+)} | B_{8l}(\mathbf{r}',\mathbf{r}'') | \Phi_1^{(+)} \Phi_l^{(-)} \rangle \langle \Phi_l^{(-)} | U_{NR}^{(1)}(\mathbf{r}''') | \Phi_2^{(-)} \rangle}{\omega_l(1-i0)-\omega-\omega_2^{(-)}} \Bigg\}. \qquad (3.5)
\end{aligned}
$$

Likewise, in (2.2) we separate out terms with intermediate terms corresponding to denominators in the resonant factors of order $2mc^2$. This part of the matrix (2.2) can be denoted by $U^{(3)}_{i\to f}\left(\begin{smallmatrix} l_- \\ l_+ \end{smallmatrix}\right)$ and written explicitly using the operator $U^{(3)}_{\text{soft}}$ (Eq. (2.12)).

Let $\Phi^{\alpha\beta}_{pr}(\mathbf{r}',\mathbf{r}'') = \Phi^{\alpha}_{p}(\mathbf{r}')\Phi^{\beta}_{r}(\mathbf{r}'')$ be the wave function of the electron–positron system, a simple product of wave functions of the individual particles in states $p,r=1,2$ $(\alpha,\beta=+,-)$. Then the matrix (3.5) can be written as a sum of terms like

$$\langle \Phi^{\alpha\beta}_{pr} | U^{(1)}_{NR} | \triangle\Phi^{\alpha\beta}_{nm} \rangle + \langle \triangle\Phi^{\alpha\beta}_{pr} | U^{(1)}_{NR} | \Phi^{\alpha\beta}_{nm} \rangle, \qquad (3.6)$$

where $\triangle\Phi^{\alpha\beta}_{nm}$ and $\triangle\Phi^{\alpha\beta}_{pr}$ are the corrections to the unperturbed functions $\Phi^{\alpha\beta}_{nm}$ and $\Phi^{\alpha\beta}_{pr}$. We next allow for the fact that matrix elements like $\langle \triangle\Phi^{\alpha\beta}_{pr} | U^{(1)}_{NR} | \triangle\Phi^{\alpha\beta}_{pr} \rangle$ are much smaller than the matrix elements (3.6). Instead of the matrix elements (3.6), we then have

$$\langle \Phi^{\alpha\beta}_{pr} + \triangle\Phi^{\alpha\beta}_{pr} | U^{(1)}_{NR} | \Phi^{\alpha\beta}_{nm} + \triangle\Phi^{\alpha\beta}_{nm} \rangle, \qquad (3.7)$$

where formally in addition to third-order effects we have included first-order effects of type $\langle \Phi^{\alpha\beta}_{pr} | U^{(1)}_{NR} | \Phi^{\alpha\beta}_{pr} \rangle$.

Actually, the introduction of the matrix elements (3.7) can be interpreted in the following way. The emission (absorption) of a real photon by a system consisting of an electron and positron interacting via the virtual photon field can be described as a first-order effect if the electron and positron are assumed to comprise a single dynamical system. This means that instead of the wave functions $\Phi^{\alpha\beta}_{pr}$ we can introduce the wave functions of the positronium atom. Depending on the choice of terms in the operators B_{sl}, we can then allow for both orbital and spin degrees of freedom of the electron and positron. Furthermore, depending on the choice of the terms in the operator $U^{(1)}_{NR}$, we can examine the different types of quantum transitions in the positronium atom.

In the problem examined in Refs. 3 and 4 — the interaction of two electrons belonging to two hydrogen-like atoms located at an arbitrary distance from each other — matrix elements of type (3.7) were also separated out in order to introduce the concept of polarization fields. It was found that first-order effects lead to a change in the quantum state of one of the electrons, which is equivalent to taking into account the external field at the point where this electron is located. Third-order effects, which do not induce quantum transitions in the spectrum of the other electron, produced a polarization field (electron or positron, depending on the type of intermediate state) due to the exchange of virtual photons. Thus, the polarization field in the problem examined in Refs. 3 and 4 is not a field of real photons, and can be considered a correction to the external field acting on the individual electrons.

The explicit form of the polarization fields in that problem depends on the type of quantum transitions (electric dipole, magnetic, etc.) in the spectrum of the interacting electrons, and corresponds to one-electron transitions. Two-electron transitions, accompanied by the emission or absorption of a real photon with twice the energy, was examined in Ref. 20. Energy conservation makes it impossible for a single atomic electron to emit or absorb a photon of twice the original energy, so first-order effects are forbidden in matrix elements of type (3.7) and the process is two-particle, i.e., it can be described as a third-order QED effect.

The electron–positron system in the positronium atom is in a similar situation. The emission and absorption of a real photon is two-particle, and can be considered a third-order effect with matrix (3.5).

On the other hand, in the nonrelativistic theory of quantum transitions, single-particle wave functions of the positronium atom can be used if we choose a reference frame with origin at the atom's center of gravity and introduce relative coordinates and momenta. The fine structure of the positronium atom can be described using appropriate electron–positron interaction operators by transforming to total spin and angular momentum operators.[13] Quantum transitions in such a system can be described as first-order perturbation effects.[10] This way of describing quantum transitions in the spectrum of the positronium atom can considered equivalent to the approach that uses the matrix (3.5) with two-particle wave functions of the positronium atom if, for example, we are only interested in calculating the probability of quantum transitions between two states of the positronium atom.

The situation changes dramatically if we need a detailed description of the kinetics of quantum transitions, as for example in the problem discussed in Ref. 12, which studies the annihilation decay of the positronium atom from two states between which spontaneous and stimulated optical transitions may occur. To describe such a process we invoke the appropriate commutation relations for fermion and boson operators and an idea about the two-particle wave functions of the positronium atom. Using single-particle wave functions would require the introduction of Pauli commutation relations. Such a change in statistics masks the physical essence of the process. Thus, the introduction of two-particle wave functions of positronium (and hence the concept of polarization fields) is productive in describing the self-consistent interaction of the positronium atom and the field of internal and external photons, in describing the gradual formation of the photon field inside the atom, etc. The explicit form of the polarization field acting on the electron or positron in the positronium atom can then be determined from the matrix (3.5) after the appropriate terms of the operator $U_{NR}^{(1)}$ (Eq. (3.2)) have been substituted into the matrix.

4. ELECTRIC DIPOLE TRANSITIONS IN THE POSITRONIUM ATOM

We now examine electric dipole transitions between states 1 and 2 of the positronium atom (for example, $1S$ and $2P$) under the following conditions:

(1) The intermediate states $l_\pm$ are chosen in such a way that the energy difference $\hbar(\omega_l \pm \omega \mp \omega_k^{\pm})$ in (2.2) is much less than $2mc^2$ $(k=1,2)$.

(2) In the operators B_{sl} we only leave terms that do not contain the speed of light, i.e., nonrelativistic terms. As shown above, such terms only appear in the soft mode of electron–positron interaction. Under such conditions the electric dipole transitions have the highest effective interaction energy, and hence the highest probability per unit time of a quantum transition involving the emission (absorption) of an optical photon.

Electric dipole transitions in a positronium atom are governed by the nonvanishing quantity

$$
\begin{aligned}
G(\mathbf{k}) = {} & -\frac{e}{mc}\sqrt{\frac{2\pi\hbar c^2}{V_R\omega}}\, e^{i\mathbf{k}\cdot\mathbf{a}}\,\frac{e^2}{4\pi\hbar}\int d\mathbf{r}'\,d\mathbf{r}''\,d\mathbf{r}''' \\
& \times \sum_{l_\pm} \left\{ \frac{-1}{\omega_l(1-i0)+\omega-\omega_2^{(+)}} \right. \\
& \times \Phi_2^{(+)*}(\mathbf{r}')(\mathbf{p}_1\cdot\mathbf{e})\Phi_l^{(+)}(\mathbf{r}')\Phi_l^{(+)*}(\mathbf{r}'') \\
& \times \Phi_1^{(-)*}(\mathbf{r}''')\frac{1}{|\mathbf{r}''-\mathbf{r}'''|}\Phi_1^{(+)}(\mathbf{r}'')\Phi_2^{(-)}(\mathbf{r}''') \\
& - \frac{1}{\omega_l(1-i0)-\omega-\omega_1^{(+)}}\Phi_2^{(+)*}(\mathbf{r}')\Phi_1^{(-)*}(\mathbf{r}''') \\
& \times \frac{1}{|\mathbf{r}'-\mathbf{r}'''|}\Phi_l^{(+)}(\mathbf{r}')\Phi_2^{(-)}(\mathbf{r}''')\Phi_l^{(+)*}(\mathbf{r}'')(\mathbf{p}_2\cdot\mathbf{e})\Phi_1^{(+)} \\
& \times (\mathbf{r}'') - \frac{1}{\omega_l(1-i0)+\omega-\omega_1^{(-)}}\Phi_1^{(-)*}(\mathbf{r}'') \\
& \times (\mathbf{p}_2\cdot\mathbf{e})\Phi_l^{(-)}(\mathbf{r}'')\Phi_2^{(+)*}(\mathbf{r}')\Phi_l^{(-)*}(\mathbf{r}''') \\
& \times \frac{1}{|\mathbf{r}'-\mathbf{r}'''|}\Phi_1^{(+)}(\mathbf{r}')\Phi_2^{(-)}(\mathbf{r}''') \\
& - \frac{1}{\omega_l(1-i0)-\omega-\omega_2^{(-)}}\Phi_2^{(+)*}(\mathbf{r}')\Phi_1^{(-)*}(\mathbf{r}'') \\
& \times \frac{1}{|\mathbf{r}'-\mathbf{r}''|}\Phi_1^{(+)}(\mathbf{r}')\Phi_l^{(-)}(\mathbf{r}'') \\
& \left. \times \Phi_l^{(-)*}(\mathbf{r}''')(\mathbf{p}_3\cdot\mathbf{e})\Phi_2^{(-)}(\mathbf{r}''') \right\}, \qquad (4.1)
\end{aligned}
$$

where V_R is the electromagnetic field quantization volume, **k** is the wave vector of an optical photon, **a** is the radius vector of the positronium atom, and **e** is the polarization unit vector of the optical photon.

To calculate $G(\mathbf{k})$ and the emission or absorption probability per unit time of a photon of frequency ω, we must know the electron and positron wave functions and the corresponding particle energy values in the positronium atom.

4.1. Electron and positron stationary states in the positronium atom

The operators (3.3) and (3.4) can be applied to the problem of positronium as a bound electron–positron system. The nonrelativistic terms in these operators then define the unperturbed Hamiltonian operator, while the relativistic terms determine the fine structure of the positronium atom (to order $1/c^2$). The unperturbed problem reduces to the hydrogen-atom problem of nonrelativistic quantum mechanics in the positronium center-of-mass system. The unperturbed Schrödinger equation will then coincide with the electron equation of motion in the hydrogen atom if the electron mass is replaced by the reduced mass $m/2$. The energies of

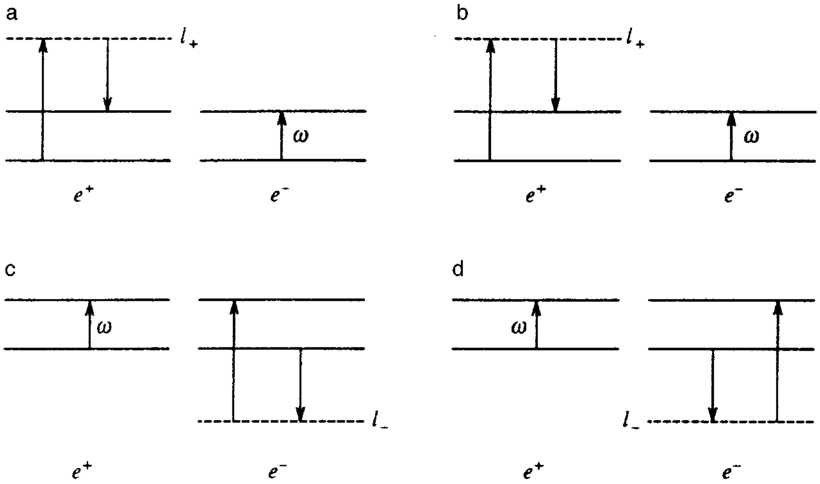


FIG. 2. Quantum transitions in the positronium atom involving absorption of a real photon.

the positronium atom are therefore half the corresponding energies of the hydrogen atom, while the radii of the orbits are twice as large.[13]

In contrast to the traditional approach of Ref. 13, we find the wave functions of the electron and positron in the positronium atom separately, as well as the set of quantum numbers of these particles. We write the unperturbed Hamiltonian operator of the positronium atom in the form

$$H=H_{01}+H_{02},$$

$$H_{01}=\frac{1}{2m}\mathbf{p}_1^2-\frac{e^2}{2|\mathbf{r}_1-\mathbf{r}_2|},\quad H_{02}=\frac{1}{2m}\mathbf{p}_2^2-\frac{e^2}{2|\mathbf{r}_1-\mathbf{r}_2|},\tag{4.2}$$

where $\mathbf{r}_1$ and $\mathbf{p}_1$ are the radius vector and momentum for the electron, and $\mathbf{r}_2$ and $\mathbf{p}_2$ are the same quantities for the positron.

To calculate the commutator $[H_{01},H_{02}]$, we must use the commutators

$$\left[\mathbf{p}_1^2,\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\right]=2i\hbar\frac{\mathbf{n}_{12}\cdot\mathbf{p}_1}{|\mathbf{r}_1-\mathbf{r}_2|^2},$$

$$\left[\mathbf{p}_2^2,\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\right]=-2i\hbar\frac{\mathbf{n}_{12}\cdot\mathbf{p}_2}{|\mathbf{r}_1-\mathbf{r}_2|^2},\tag{4.3}$$

where $\mathbf{n}_{12}=(\mathbf{r}_1-\mathbf{r}_2)/|\mathbf{r}_1-\mathbf{r}_2|$. Allowing for the fact that $\mathbf{p}_1=-\mathbf{p}_2$, in the positronium atom, we find that $[H_{01},H_{02}]=0$. This means that we can introduce the constants of the motion l_1^2, l_{1z}, and H_{01} for the electron, and l_2^2, l_{2z}, and H_{02} for the positron, with the corresponding set of quantum numbers l_1, m_{1z}, n_1 and l_2, m_{2z}, n_2. Following Ref. 21, for particle motion in a central field, we find the wave functions and energies of the electron:

$$\Phi^{(+)}_{n_1l_1m_1}=R_{n_1l_1}(r)Y_{l_1m_1}(\Theta,\varphi),\quad E_{n_1}=-\frac{1}{2n_1^2}\frac{me^4}{4\hbar^2},\tag{4.4}$$

where $n_1\geqslant l_1+1$ is the principal quantum number, l_1 is the orbital quantum number, and m_1 is the magnetic quantum number of the electron. The radial wave functions can be defined in terms of Laguerre polynomials,[21] and show that the radii of the electron orbits in the positronium atom is twice the radii of the electron orbits in the hydrogen atom. The electron energy values E_{n_1} are one fourth the values in the hydrogen atom. For the positron in the positronium atom we obtain similar results.

According to Ref. 13, the energy of the positronium atom is

$$E_n=-\frac{1}{4}\frac{me^4}{n^2\hbar^2},\tag{4.5}$$

where n is the principal quantum number. Comparing (4.5) and (4.4), we find that $n_1=n_2=1$ when $n=1$, $n_1=n_2=2$ when $n=2$, etc. The orbital and magnetic quantum numbers of the positronium atom, L and M, can be expressed in terms of the quantum numbers l_1, m_1 and l_2, m_2 if we examine the problem of adding the angular momenta $\mathbf{l}_1$ and $\mathbf{l}_2$ of the electron and positron.[21] Here we have $M=m_1+m_2$ and

$$L=l_1+l_2,l_1+l_2-1,\dots,|l_1-l_2|.\tag{4.6}$$

This means that in the positronium atom, the principal quantum numbers of positronium, the electron, and the positron coincide, while the orbital and magnetic quantum numbers may be different. Thus, in the $1S$ state of the positronium atom we have

$$n=n_1=n_2=1,\quad m_1=m_2=0,\quad L=0,\quad l_1=l_2=0.$$

In the $2P$ state we have

$$n_1=n_2=n=2,\quad L=1,$$

$$l_1=0,\quad l_2=1,\quad m_1=0,\quad m_2=0,\pm1$$

or

$$l_1=1,\quad l_2=0,\quad m_2=0,\quad m_1=0,\pm1.$$

The wave functions of positronium can be expressed as a linear combination of simple products of electron and positron wave functions with appropriate Clebsch–Gordan coefficients.[20]

4.2. Polarization fields in the positronium atom

Figure 2 depicts the various quantum transition diagrams corresponding to each term in (4.1). Let us examine the first diagram for the absorption of an optical photon. According to the expansion of the general solution (2.2a) of the Dirac equation, for this diagram we have the initial state of the electron and positron with the wave function $\Phi_1^{(+)} \times(\mathbf{r}'')\Phi_1^{(-)*}(\mathbf{r}''')$ and the energies $2E_1$, with E_1 specified in (4.1). As a result of the electron–positron interaction via the field of virtual photons, the electron finds itself in the intermediate state l_+ with positive energy, and the positron in state 2. Then a real photon is absorbed and the electron goes from state l_+ to state 2. The second diagram representing the absorption of an optical photon, Fig. 2b, differs from the first. Indeed, first there is a transition of the electron from state 1 to state l_+ initiated by a real photon of frequency ω. Then, as a result of the electron–positron interaction via the field of virtual photons, the electron and positron go into state 2. The diagrams in Figs. 2c and d are similar to the first two, but the transitions involve intermediate states with negative energy $\hbar\omega_{l_-}$.

Using the expression (4.1), we can write explicit equations for the vector potentials of the fields at points in space where the electron and positron are located, with the radius vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Here we allow for the multimode nature of particle interaction and introduce the polarization vectors $\mathbf{e}_{k\lambda}$, the frequencies ω_k, and the wave vectors $\mathbf{k}$ of the various photons participating in the induction of polarization fields ($\lambda=1,2$ are the photon polarization indices). To this end, we write the interaction of the electron and the polarization field $\mathbf{A}_e(\mathbf{r}_1,t)$ in terms of the operator $U_{NR1}^{(1)}=-(e/mc)\mathbf{p}_1\cdot\mathbf{A}_e$, where

$$\begin{aligned}
\mathbf{A}_e(\mathbf{r}_1,t)=\sum_{k\lambda}\ &\sqrt{\frac{2\pi\hbar c^2}{V_R\omega_k}}e^{-i\omega_k t}\mathbf{e}_{k\lambda}c_{k\lambda}\lambda e^{ikr_1}\\
&\times\frac{1}{\langle\Phi_2^{(+)}|\mathbf{p}_1\cdot\mathbf{e}_{k\lambda}|\Phi_1^{(+)}\rangle}\\
&\times\sum_{l_+}\left\{\frac{-1}{\hbar[\omega_l(1-i0)+\omega_k-\omega_2^{(+)}]}\right.\\
&\times\langle\Phi_2^{(+)}|\mathbf{p}_1\cdot\mathbf{e}_{k\lambda}|\Phi_l^{(+)}\rangle\\
&\times\left\langle\Phi_l^{(+)}\Phi_l^{(-)}\left|\frac{e^2}{4\pi|\mathbf{r}''-\mathbf{r}'''|}\right|\Phi_1^{(+)}\Phi_2^{(-)}\right\rangle\\
&-\frac{1}{\hbar[\omega_l(1-i0)-\omega_k-\omega_1^{(+)}]}\\
&\times\left\langle\Phi_2^{(+)}\Phi_1^{(-)}\left|\frac{e^2}{4\pi|\mathbf{r}'-\mathbf{r}'''|}\right|\Phi_l^{(+)}\Phi_2^{(-)}\right\rangle\\
&\left.\times\langle\Phi_l^{(+)}|\mathbf{p}_2\cdot\mathbf{e}_{k\lambda}|\Phi_1^{(+)}\rangle\right\}.
\end{aligned}\tag{4.1a}$$

Following the same line of reasoning, we can write the vector potential $\mathbf{A}_p(\mathbf{r}_2,t)$ of the polarization field acting on the positron in the positronium atom by using the operator $U_{NR2}^{(1)}=-(e/mc)\mathbf{p}_2\cdot\mathbf{A}_p$, where

$$\begin{aligned}
\mathbf{A}_p(\mathbf{r}_2,t)=\sum_{k\lambda}\ &\sqrt{\frac{2\pi\hbar c^2}{V_R\omega_k}}e^{-i\omega_k t}\mathbf{e}_{k\lambda}c_{k\lambda}\lambda e^{ikr_2}\\
&\times\frac{1}{\langle\Phi_2^{(-)}|\mathbf{p}_2\cdot\mathbf{e}_{k\lambda}|\Phi_1^{(-)}\rangle}\\
&\times\sum_{l_+}\left\{\frac{-1}{\hbar[\omega_l(1-i0)+\omega_k-\omega_1^{(-)}]}\right.\\
&\times\langle\Phi_1^{(-)}|\mathbf{p}_2\cdot\mathbf{e}_{k\lambda}|\Phi_l^{(-)}\rangle\\
&\times\left\langle\Phi_2^{(+)}\Phi_l^{(-)}\left|\frac{e^2}{4\pi|\mathbf{r}'-\mathbf{r}'''|}\right|\Phi_1^{(+)}\Phi_2^{(-)}\right\rangle\\
&-\frac{1}{\hbar[\omega_l(1-i0)-\omega_k-\omega_1^{(-)}]}\\
&\times\left\langle\Phi_2^{(+)}\Phi_1^{(-)}\left|\frac{e^2}{4\pi|\mathbf{r}'-\mathbf{r}''|}\right|\Phi_1^{(+)}\Phi_l^{(-)}\right\rangle\\
&\left.\times\langle\Phi_l^{(-)}|\mathbf{p}_2\cdot\mathbf{e}_{k\lambda}|\Phi_2^{(-)}\rangle\right\}.
\end{aligned}\tag{4.1b}$$

Obviously, polarization fields also emerge when optical photons are emitted. In this case we must replace ω_k in (4.1a) and (4.1b) by $-\omega_k$, $\mathbf{k}$ by $-\mathbf{k}$, and the operator $c_{k\lambda}$ by the creation operator $c_{k\lambda}^{\dagger}$ for a photon of mode $\mathbf{k}\lambda$, and also interchange the electron and positron states, $1\leftrightarrow 2$.

Equations (4.1a) and (4.1b) make it possible to study the spatial–temporal behavior of the fields in the positronium atom, having chosen the states 1 and 2 of the electron and positron, including Rydberg states. The number of photon modes participating in field formation is determined by the energy conservation law (2.3). It must be noted, however, that in the early stages of field formation the conservation law is not specified by a delta function, so that the number of photon modes can be large. A detailed analysis of the fields (4.1a) and (4.1b) as functions of coordinates and time requires further investigation.

4.3 Probability of spontaneous photon emission per unit time

We now apply Eq. (4.1) to the spontaneous emission of a single real photon. We first replace ω with $-\omega$ in (4.1) and permute the indices: $1\rightleftarrows 2$. In accordance with the conservation law (2.3), for the $2P\rightarrow 1S$ transition we have

$$\hbar\omega=\frac{3}{16}\frac{me^4}{\hbar^2}.\tag{4.7}$$

The wavelength λ corresponding to this transition is 2465Å. According to Ref. 9, the lifetime τ of the $2P$ state of positronium is 3.2×10^{-9} s. Indeed, this result can easily be obtained if we use single-particle wave functions of the positronium atom[10] and the energy values (4.5) to calculate τ. Accordingly, the probability of spontaneous photon emission per unit time for this transition is $W=\tau^{-1}=0.3125\times 10^9\ \mathrm{s}^{-1}$.

We now use two-particle wave functions of positronium (see Sec. 4.1) to calculate $G(\mathbf{k})$ and W. As Eq. (4.1) shows, we need to calculate the matrix elements of the electron or positron momentum operator and the matrix elements of the electron–positron Coulomb interaction operator, allowing for various intermediate states $l_\pm$. To this end, it is convenient to expand the Coulomb interaction operator in the matrix elements in terms of spherical harmonics $Y_{kq}(\Theta_1,\varphi_1)$ and $Y_{kq}(\Theta_2,\varphi_2)$ (see Ref. 21):

$$\frac{e^2}{r_{12}}=e^2\sum_{k=0}^{\infty}\frac{4\pi}{2k+1}\frac{r_<^k}{r_>^{k+1}}\sum_q Y_{kq}(\Theta_1,\varphi_1)Y_{kq}^*(\Theta_2,\varphi_2), \tag{4.8}$$

where Θ_1, φ_1, Θ_2, and φ_2 are the angular variables of the electron and positron, and $r_<$ and $r_>$ denote the lesser and greater of the lengths of the radius vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. Following the method of calculating matrix elements in atomic spectroscopy,[21] we find that calculations of the probability W of the optical transition based on formula (4.1) based on two-particle wave functions yield the same result as in Refs. 9 and 10.

5. CONCLUSION

In this paper we have given a detailed substantiation of the physical meaning of the coupling constant $G(\mathbf{k})$ in the effective Hamiltonian (2.1). For electric dipole transitions, we have derived a formula (Eq. (4.1)) that allows for various types of electron–positron interaction in the field of optical and annihilation photons. For other quantum transition types, with allowance for relativistic terms in the operator B_{sl}, Eq. (4.1) can easily be generalized by using the matrix (2.2) and analyzing the situation.

We have solved the problem of electron–positron interaction accompanied by emission or absorption of a soft photon. Such interaction can be interpreted as a third-order QED effect with the Feynman diagrams of Fig. 1. Here are the main properties of this interaction.

A key feature of the diagrams in Fig. 1 is the exchange of soft and hard photons. In accordance with this, we can separate the soft and hard modes of the electron–positron interaction, which are described by the operators B_{sl}. If there is no emission (absorption) of a real photon, the electron–positron interaction is a second-order QED effect. In this case, the operators B_{sl} cease to depend on the intermediate states $\omega_{l_\pm}$ and the soft interaction mode corresponds to the operator $B_1 \equiv U^{(2)}$, where $U^{(2)}$ is the Breit operator. The hard interaction mode is described by the operator B_5 (Eq. (3.4)).

When emission (absorption) of a real photon is taken into account in the Feynman diagrams of Fig. 1, the exchange of virtual photons leads to the emergence of intermediate states. As a result, the generalized operators B_{sl} begin to depend on the frequencies $\omega_{l_\pm}$ of the intermediate states.

Since the electron and positron masses are equal, we are able to deal with two-particle wave functions of the positronium atom. This makes it possible to introduce the concept of polarization fields inside the positronium atom. We believe that this concept plays an important role in precision optical and radio spectroscopy of the positronium atom and in quantum positronics.

The work was supported by the Russian Fund for Fundamental Research.

[1] M. Born and E. Wolf, *Principles of Optics*, 6th ed., Pergamon Press, Oxford (1980).

[2] O. N. Gadomskiĭ, V. P. Nagibarov, and N. K. Solovarov, Zh. Éksp. Teor. Fiz. **63**, 813 (1972) [Sov. Phys. JETP **36**, 426 (1973)].

[3] O. N. Gadomskiĭ, V. P. Nagibarov, and N. K. Solovarov, Zh. Éksp. Teor. Fiz. **70**, 435 (1976) [Sov. Phys. JETP **43**, 225 (1976)].

[4] O. N. Gadomskiĭ and R. A. Vlasov, *Optical Echo Spectroscopy of Surfaces* [in Russian], Nauka and Tekhnika, Minsk (1990).

[5] O. N. Gadomsky and K. V. Krutitsky, Quantum Semiclassic. Opt. **9**, 343 (1997).

[6] O. N. Gadomskiĭ and K. V. Krutitskiĭ, Zh. Éksp. Teor. Fiz. **106**, 936 (1994) [JETP **79**, 513 (1994)].

[7] O. N. Gadomsky and K. V. Krutitsky, J. Opt. Soc. Am. B **13**, 1679 (1996).

[8] G. W. F. Drake, Phys. Rev. **5**, 1979 (1972).

[9] K. P. Ziock, C. D. Dermer, R. H. Howell, and F. Magnotta, Preprint UCRL-100160 (1989).

[10] C. D. Dermer and J. C. Weisheit, Phys. Rev. A **40**, 5526 (1989).

[11] A. Rich, Rev. Mod. Phys. **53**, 127 (1981).

[12] O. N. Gadomskiĭ, Zh. Éksp. Teor. Fiz. **110**, 1228 (1996) [JETP Lett. **83**, 676 (1996)].

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

[14] E. M. Lifshitz, Zh. Éksp. Teor. Fiz. **18**, 562 (1948).

[15] B. N. Fedushin, Zh. Éksp. Teor. Fiz. **22**, 140 (1952).

[16] G. M. Garibyan, Zh. Éksp. Teor. Fiz. **24**, 617 (1953).

[17] J. Pirenne, Arch. Sci. Phys. Nat. **29**, 207 (1947).

[18] V. B. Berestetskiĭ and L. D. Landau, Zh. Éksp. Teor. Fiz. **19**, 673 (1949).

[19] C. S. Chang and P. Stehle, Phys. Rev. **4**, 630 (1971).

[20] O. N. Gadomskiĭ, Teor. Mat. Fiz. **106**, 145 (1996).

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

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Effect of the initial charge and charge-state fluctuations on the range parameters of high-energy ions

F. F. Komarov and A. F. Komarov

A. N. Sevchenko Research Institute of Applied Physics, Belorussian State University, 220064 Minsk, Republic of Belarus

(Submitted 10 March 1997)
Zh. Éksp. Teor. Fiz. **113**, 489–498 (February 1998)

A model used in calculating the depth distribution of implanted atoms, and the energy released into the electron and nuclear subsystems of the solid is studied. It takes into account the initial charge and the fluctuations of charge states of high-energy ions with energies $E \geqslant 1$ MeV amu^{-1}. The results of calculations are obtained on the basis of solutions of transport equations. Satisfactory agreement between theoretical predictions and experimental results is observed for a broad class of ion–target systems. [S1063-7761(98)00702-1]

1. INTRODUCTION

The penetration of ions through solids and the major processes accompanying such penetration (defect formation, atomization of the target) of medium-energy ions ($E \leqslant 10$ MeV) are described well by the Monte Carlo method or by solving the direct and inverse Boltzmann kinetic equation. However, recent experimental results[1–5] obtained for high-energy ion implantation ($E \geqslant 1$ MeV amu^{-1}) differ considerably from theoretical predictions, both for the distributions of implanted atoms and for the initial radiation-induced defects that are formed in the process. For a number of ion–target combinations the experimental values of the mean straggling of projected ranges, ΔR_p, exceed the theoretical values by a factor of 1.5 to 5, and the experimentally measured asymmetries of the depth distribution profiles of the implanted atoms are strongly shifted in the positive direction in comparison to the calculated profiles.

According to our calculations (Ref. 6), these discrepancies between theory and experiment cannot be explained by the crystallographic channeling effect, since the difference in the shapes of the ion-range distribution functions is observed near the mean projected range R_p. We believe[7–9] that the observed effects can be explained by the following reasoning. In standard theoretical approaches to the problem, the changes in the charge state of ions (electron capture and stripping) are not considered explicitly. The mean effective ion charge (Z_{eff}), which is determined by the dynamic equilibrium between electron capture and loss and depends solely on ion energy, is used as a parameter in calculating the ion scattering and stopping cross sections. Hence the effect of fluctuations in the charge states of ions on implantation processes is not taken into account in this setting.

Such an approximation works only in the low- and medium-energy ranges because of the large cross sections of electron capture and loss by ions. At energies of about 1 MeV amu^{-1} and higher the situation changes dramatically. In the energy range considered here the cross sections of electron capture and loss are so low that the characteristic ion range under conditions of unchanged charge states becomes comparable to the straggling for the profiles of implanted atoms and radiation-induced defects. The electron loss and capture cross sections $\sigma_{\pm}$ at such energies may be 10^{-4} nm^2 or even lower. Hence the probability of charge-exchange processes, $P = N\sigma_{\pm}x$ becomes close to unity at distances of order 1 μm from the target surface. Calculations done by Burenkov *et al.*[9] show that the equilibrium charge distribution for the ion flux is formed only after the ions have travelled a certain distance inside the target: $z \approx 0.8$ μm for 50-MeV boron ions B$^+$ in silicon, $z = 0.9$ μm for 60-MeV nitrogen ions Ni$^+$ in diamond, $z = 1$ μm for 59-MeV nickel ions Ni$^+$ in diamond, and $z = 7.2$ μm for 130-MeV xenon ions Xe$^+$ in diamond. Thus, random fluctuations in the charge of high-energy ions can strongly affect the shape of the distributions of the implanted atoms and radiation-induced defects. Note that these electron stopping fluctuations differ in their nature from those studied by Bohr,[10] which account for the broadening of the energy spectrum of the ions because of the random distribution of impact parameters in collisions.

In recent years there has been an upsurge of interest in the problem of theoretically describing the loss of energy by fast ions with allowance for charge-exchange processes[11–13] in connection with measurements of dE/dx in layers of materials when the layer thickness varies for various initial charge states.[14,15]

The present paper is a development of an approach suggested in Refs. 7–9. Here we allow for the effect of fluctuations in the charge state of ions not only on the ion-range characteristics but also on the depth distributions of the energy released into the electron and nuclear subsystems of the solid. We also discuss the effect of the initial charge state of the ions at the entrance to the target on the range parameters. The role that the atomic number of the target ions plays in the manifestation and the relative value of the above effects is examined. For the numerical calculations of the characteristics of implantation profiles we used the solutions of the direct and inverse (the method of distribution moments) Boltzmann equations.

2. THEORETICAL MODEL

In our case the distribution function of the ion flux in the target, $\Phi(x;\eta,E,q)$, must depend not only on the initial ion energy, the depth x in the target, the angle θ between the direction normal to the surface and the direction of ion propagation ($\eta=\cos\theta$) but also on the ion charge q. The direct kinetic equation for the distribution function of the ion flux is

$$\begin{aligned}\eta\frac{\partial\Phi(E,\eta,q;x)}{\partial z}=&N\left[\int d\sigma_n'\Phi(E',\eta',q;x)\right.\\&\left.-\int d\sigma_n\Phi(E,\eta,q;x)\right]\\&+N\frac{\partial}{\partial E}\left[\Phi(E,\eta,q;x)S_e(E,q)\right]\\&+N\sum_{q'}\left[\sigma_{q'\to q}\Phi(E,\eta,q';x)\right.\\&\left.-\sigma_{q\to q'}\Phi(E,\eta,q';x)\right],\qquad(1)\end{aligned}$$

where $d\sigma_n$ is the differential scattering cross section for an ion, with scattering amounting to the transition of the ion from a state with energy E, direction of motion η, and charge q to a state with parameters E', η', and q, respectively; $d\sigma_n'$ is the differential cross section of the inverse transition; S_e is the electron stopping cross section; and N is the concentration of atoms in the target.

In Eq. (1) the last (collision) term directly describes fluctuations of the ion charge states. Summation over q' is done over all possible charge states.

Assuming that processes such as scattering and change of the charge state of a particle are mutually independent, we cab write the inverse kinetic equation for this particle as

$$\begin{aligned}\eta\frac{\partial\Phi(r,n,E,q)}{\partial r}=&N\int\left[d\sigma_n\Phi(r,n,E,q)-\Phi\right.\\&\left.\times(r,n',E-T,q)\right]\\&+NS_e(E,q)\frac{\partial\Phi(r,n,E,q)}{\partial E}\\&-\frac{1}{2}N\Omega_e^2(E,q)\frac{\partial^2\Phi(r,n,E,q)}{\partial E^2}\\&+N\sum_{q'}\sigma_{q\to q'}\left[\Phi(r,n,E,q)\right.\\&\left.-\Phi(r,n,E,q')\right].\qquad(2)\end{aligned}$$

Here we assume that $d\sigma=d\sigma_n+d\sigma_e$. At high energies the contribution of electron stopping to the total energy losses are dominant. Hence the effect of charge fluctuations on the elastic scattering in this model can be ignored. In Eq. (2), $d\sigma_n$ is the differential elastic scattering cross section, which characterizes the transition of an ion from a state with energy E, direction of motion n, and charge q to a state with parameters $E-T$, n', and q, respectively, with T the energy transferred from the ion to the target in a single elastic collision; $\Omega_e^2(E,q)$ is the mean square straggling of electron stopping.

In contrast to the standard approaches of Refs. 16–18, to allow for fluctuations in electron stopping we must take into account the dependence of electron stopping on the ion charge, given by

$$S_e(E,q)=Z_{\mathrm{eff}}^2(E,q)S_p(E),\qquad(3)$$

where $Z_{\mathrm{eff}}(E,q)$ is the effective charge, which depends on the degree q of ionization of the ions; and $S_p(E)$ is the cross section of electron stopping for protons. Assuming that the equipartition rule is valid in the Lindhard energy loss from near and distant collisions, we used the following formula to calculate the effective charge:

$$Z_{\mathrm{eff}}^2(E,q)=\frac{1}{2}Z_{\mathrm{eff}}^2(E)+\left[1+\left(\frac{q}{Z_{\mathrm{eff}}(E)}\right)^2\right],\qquad(4)$$

where $Z_{\mathrm{eff}}(E)$ is the mean effective charge of the ions at a given energy E, calculated in the Brandt–Kitagawa model.[18]

In this case the electron stopping cross section is a statistical quantity whose distribution parameters are described in terms of cross section of the variation of the charge state. The electron capture and stripping cross section were calculated on the assumption that the charge distributions are Gaussian, which is in good agreement with the experimental data.[19] The mean values q_0 of the ion charge were calculated on the basis of the model of Ziegler *et al.*,[16] while the variance of the ion distribution over the charges was represented by the simple Bohr formula[10]

$$d=0.5\left[q_0\left(1-\frac{q_0}{Z}\right)^{5/3}\right]^{1/2},\qquad(5)$$

where Z is the nuclear charge of the ion. With the Gaussian distribution of the charge states and the dominance of single-particle processes in the variation of the charge states of ions, the cross section of variation of the charges of the ions moving in the target is described by the following expression:[19]

$$\sigma_{q\to q\pm1}=\sigma_0\exp[\mp\alpha(q-q_0)],\qquad(6)$$

where $\sigma_{q\to q\pm1}$ is the cross section of loss $(+)$ or capture $(-)$ of a single electron by a moving ion, and $\alpha=0.5d^{-2}$. For ion energies $E<E_c=50Z_T^{1/3}q_0^{1/2}$ keV amu^{-1} we use the Bohr formula[10]

$$\sigma_0=\pi a_0^2Z_T^{1/3}q_0^2\left(\frac{v_0}{v}\right)^3,\qquad(7)$$

while for $E>E_c$ we use the formula proposed by Nikolaev *et al.* (see Ref. 19):

$$\sigma_0=2\pi a_0^2Z_T^{2/3}q_0^{5/2}\left(\frac{v_0}{v}\right)^5.\qquad(8)$$

In (7) and (8), a_0 and v_0 are the atomic units of length and velocity, v is the ion velocity, and Z_T is the nuclear charge of the target atoms.

The cross section $\sigma_{q\rightarrow q+1}$ can be obtained as a function of the ion velocity and the target's atomic number if we equate the model cross section (6) and the actual electron-loss cross section calculated by formulas (7) and (8) at $q=q_0$. In other words, the fact that $\sigma_0=\sigma_{q\rightarrow -1}$ at $q=q_0$ is used to determine σ_0.

3. LIMITS OF THE METHOD

The adopted model makes it possible to take into account the initial charge of the ions and the fluctuations of the charge state of the ions by using the solution of Boltzmann kinetic equations, but requires knowing the electron loss and capture cross sections. Rigorous calculations of the electron loss and capture cross sections can be done only for low-Z ions. For medium-mass and heavy ions the approximate methods of Bohr[10] and Nikolaev *et al.*[19] are used in the theoretical treatment of this problem, while the experimental data can be found, for instance, in the review article by Betz.[19] The cross sections obtained in the present work (see Eqs. (5)–(8)) agree to within 20–30% with the experimental data.[19] This leads to a 25–40% uncertainty in calculating R_p, ΔR_p and Sk (Sk is the asymmetry of the profile of the distribution of implanted atoms over the depth of the crystal), which is severalfold smaller than the role of the effect of fluctuations of charge states of ions discussed above. However, at relatively low energies (tens of keV amu^{-1}), the calculated cross sections were found to exceed the experimental values by a factor of two to three. To such ion energies there correspond finite sections of the ion ranges, for which elastic processes of interaction of the ions and the atoms of the stopping medium become dominant. As a result of selecting the formula of Ziegler *et al.*[16] for the electron stopping cross section instead of the more common Bethe formula,[20] the ion energy range is not limited entirely to that of the first Born approximation, i.e., it extends considerably into the range of low ion energies ($v \geqslant v_F$, where v is the ion energy, and v_F is the Fermi-electron energy).

Calculations were done using a mesh grid with a depth-variable mesh width. The resulting depth distributions proved to have low sensitivity to the mesh width. Equation (1) was solved by a step-by-step calculation of the distribution function in the group approximation. For medium-mass ions (such as, say, Al$^+$), the energy and angle ranges were split into groups (20 to 30 groups in energy and 5 to 15 groups in angles). For light ions (such as, say, B$^+$) and high energies ($E>1$ MeV amu^{-1}), we were forced to increase the number of groups in energy to 50–100 so that our results would be in good agreement with the experimental data of Ref. 1, for example. In solving Eq. (2) numerically, we varied the number of steps from 100 to 1000, depending on the width of the energy interval and ion mass (the number of partitions must be increased as the ion mass decreases). The error in numerical calculations of R_p, ΔR_p, and Sk amounted to roughly 1–3%.

4. RESULTS OF CALCULATIONS AND DISCUSSION

To evaluate the effect of charge distribution fluctuations on the formation of the distribution profiles of implanted atoms, we calculated the parameters of the implantation profiles (R_p, ΔR_p, and Sk) with and without allowance for ion charge fluctuations. Calculations with allowance for charge distribution fluctuations were done using the method of moments to numerically solve the system of integro-differential equations (2). The cross sections of elastic scattering and electron stopping were take from Ref. 17, while the cross sections of electron capture and loss by ions were calculated by Eqs. (5)–(8). The method of numerically solving the inverse kinetic equation is described in Ref. 17. The difference between Eq. (2) and the standard inverse transport equation of Ref. 17 lies in the presence of a collision term in the former, a term that describes the change in the states of the ion charges. This difference does not preclude using the procedure of numerical solution of equations of this type developed earlier. Nevertheless, the structure of the system of linear equations approximating (2) is somewhat more complicated than in the common case discussed in Ref. 17.

In contrast to the ordinary approach, our method makes it possible to model the implantation of ions with various initial charge states. The parameters of the profiles for implantation of boron and aluminum atoms into silicon are listed in Tables I and II as functions of the initial state of the ions. As expected, the mean projected ion range R_p increases as the initial charge of the ions decreases (in the case of boron ions the increase may amount to 6%). This occurs because of the decrease in electron stopping at small depths, where the equilibrium distribution of atoms over the charge states has not enough time to set in. More important, however, is the effect of the initial charge states on the higher-order moments of the distributions of the implanted atoms (ΔR_p and Sk). For instance, the discrepancies in the values of ΔR_p amount to 23% and 12% for the implantation of boron and aluminum atoms, respectively. For the distribution asymmetry factor Sk, the maximum discrepancy amounts to 48% and 27% for the implantation of boron and aluminum

TABLE I. Spatial moments of the distribution of boron with energy 100 MeV implanted in silicon (with allowance for fluctuations in ionic charge) for various initial charge states of the incident ions.

q	R_p, μm	ΔR_p, μm	Sk
1	313.9	11.78	0.49
2	311.3	11.82	0.50
3	307.1	11.54	0.52
4	301.2	10.65	0.61
5	295.2	9.43	0.78

TABLE II. Spatial moments of the distribution of aluminum with energy 100 MeV implanted in silicon (with allowance for fluctuations in ionic charge) for various initial charge states of the incident ions.

q	R_p, μm	ΔR_p, μm	Sk
1	40.40	0.97	−3.59
2	40.37	0.96	−3.60
6	40.16	0.95	−3.73
8	39.98	0.93	−3.94
12	39.50	0.86	−4.70

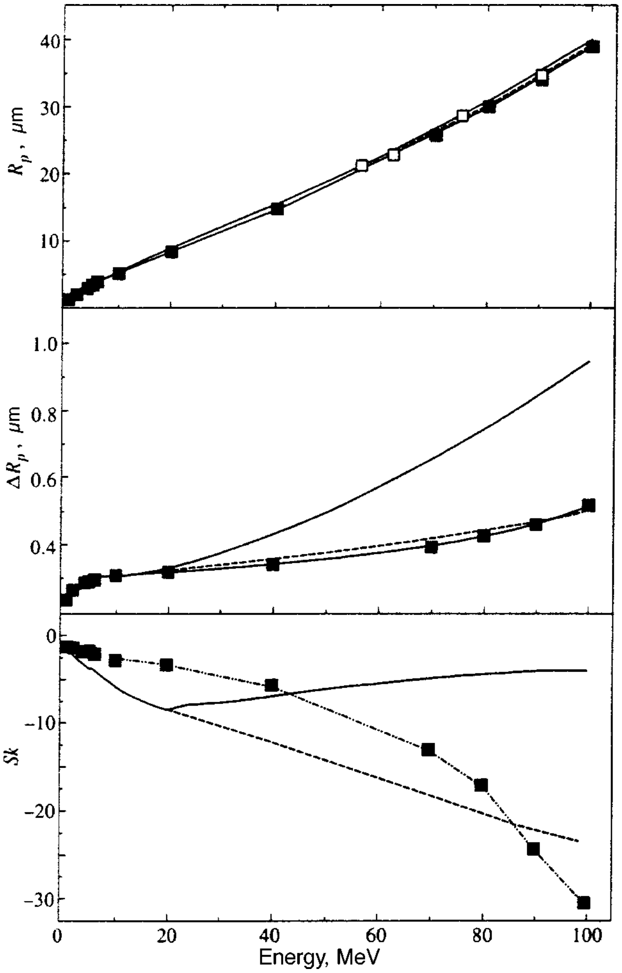


FIG. 1. Spatial moments of the distribution of aluminum in silicon at implantation energies up to 100 MeV. Solid curves depict the results of calculations obtained with allowance for fluctuations in the charge states of the ions, dashed curves depict the results of calculations without allowance for charge fluctuations,[17] the ■ indicate the results of standard modeling,[16] and the □ indicate the experimental data of La Ferla *et al.*[2].

atoms, respectively. As the ion energy rises, the effect of the initial charge state of the ion on the parameters of the implantation profiles gets stronger. The results of these calculations make it possible to conclude that the strong discrepancy between the distribution moments of the implanted atoms obtained in the given model and those obtained on the basis of standard calculations[16,17] is due to the effect of ion charge fluctuations rather than to the choice of the initial charge state of the implanted ions. Experiments in which postimplantation profiles are measured suggest that the fluctuations of the charge state of the ions have a strong effect of the formation of impurity and defect profiles when high-energy ion implantation is involved.

In Fig. 1 we depict, for the sake of comparison, the experimental and theoretical values of the parameters of the distribution of atoms over the target depth as functions of the ion energy for the case in which silicon is bombarded by aluminum atoms. The calculations were done with and without allowance for fluctuations of the ion charge distributions. Allowing for charge state fluctuations modifies the value of R_p very slightly. At the same, such fluctuations have a strong effect on ΔR_p and especially on Sk. For instance, in the case of implantation of aluminum atoms in silicon at an energy $E = 100$ MeV, the value of ΔR_p nearly doubles. The variation of Sk may be considerably greater. As the ion mass increases, the difference in the values of the ion-range parameters show up at high energies, for which broader charge distributions of the ion flux are formed in a sizable segment of the ion range in the target.

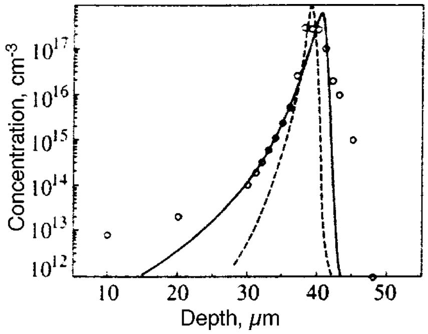


FIG. 2. Distribution of aluminum with energy 100 MeV implanted in silicon. The solid curve depicts the calculated profile obtained with allowance for fluctuations in charge state of the ions, the dashed curve represents the results of standard modeling,[16] and the ○ indicate the experimental data of La Ferla *et al.*[2].

Numerical solution of the direct kinetic equation (1) makes it possible to extract information about the total depth distributions of the implanted atoms (Fig. 2) and the distribution profiles of the energy released in inelastic (track formation) and other (defect formation) processes (Fig. 3a). Figure 2 shows that the results of calculations with our model correspond better to the experimental data than those of standard calculations[16] for the distribution profile of aluminum atoms implanted in silicon at $E = 100$ MeV. The experimental profiles exhibit a greater dispersion of ions over the ranges than the one yielded by our model, which is an indication of the highly approximate nature of the ion charge-exchange cross section used in the present paper. The considerably higher dispersion of ions over the range in our model (compared to that obtained by standard calculations) is also characteristic of depth distributions of radiation-induced defects created as a result of elastic collisions with target atoms (curve *4* in Fig. 3a).

Figure 3b depicts the calculated distribution profiles for 250-MeV xenon atoms implanted in indium phosphide obtained in our model (curve *2*) and in standard calculations[16] (curve *1*).

It is hardly possible to overestimate the importance of information about the distributions of the energy released by ions elastically (curve *4* in Fig. 3a) and in collisions with electrons of the stopping medium (curve *2* in Fig. 3a). First, using the dE/dx profile, one can calculate the dose of high-energy ions at which a latent amorphous layer is formed deep inside semiconductors and insulators near the maximum in the elastically released energy. For instance, in view of the energy criterion, amorphization of a local volume of a semiconducting crystal occurs when the energy elastically released within a unit volume reaches a critical value $\varepsilon_{\mathrm{cr}}$. For different semiconductors $\varepsilon_{\mathrm{cr}}$ varies over a broad range.

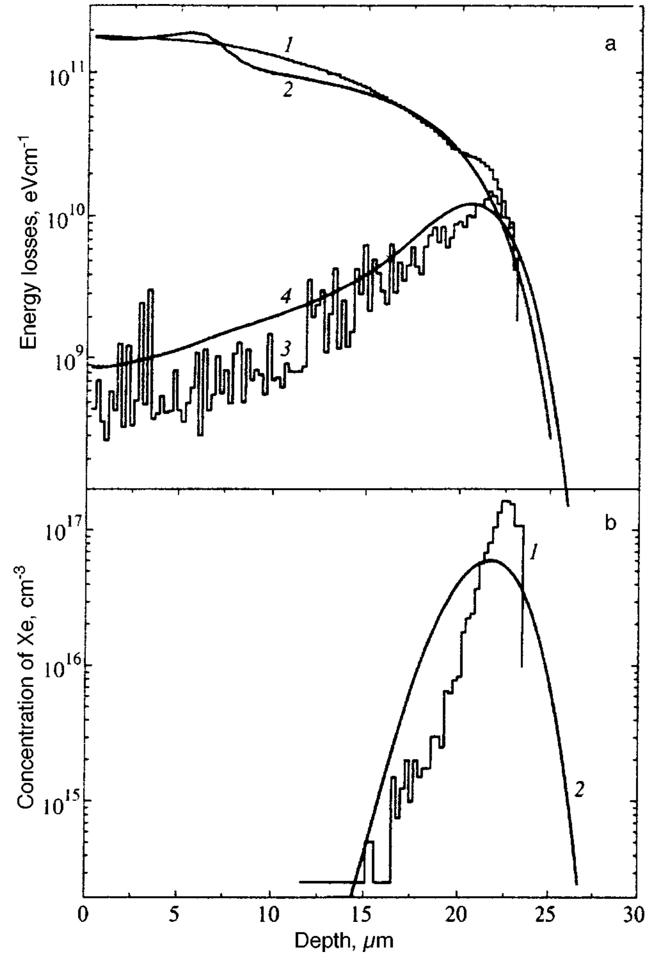


FIG. 3. (a) Distribution profiles for the energy liberated in inelastic (curves *1* and *2*) and elastic (curves *3* and *4*) processes as a result of implanting xenon with energy 250 MeV in indium phosphide. Curves *1* and *3* represent the results of standard modeling[16] and curves *2* and *4*, the results obtained on the basis of the present model. (b) Distribution profiles for xenon with energy 250 MeV implanted in indium phosphide calculated on the basis of the present model (curve *2*) and the standard model[16] (curve *1*).

For instance, for silicon $\varepsilon_{\mathrm{cr}}=6\times10^{20}$ keV/cm^3 at lower temperatures and increases to 5×10^{21} keV/cm^3 at room temperature.[20] For GaAs crystals, the characteristic values of $\varepsilon_{\mathrm{cr}}$ are $(2.5-3)\times10^{20}$ keV/cm^3 at low temperatures and $(8.3-3)\times10^{20}$ keV/cm^3 at room temperature,[21,22] while for InP this value is smaller by a factor of 5 to 10.

Since

$$\varepsilon_{\mathrm{cr}}=NS_n^{\max}D, \tag{9}$$

where N is the concentration of atoms in the crystal, $S_n=(1/N)(dE/dx)_n$, and D is the ion dose, we can easily calculate the ion dose needed for the formation of a latent amorphous layer. For example, for $E=250$-MeV xenon ions Xe$^+$ in InP (Fig. 3), the amorphization ion dose D_{amo} is approximately $(2-5)\times10^{12}$ ion/cm^2 at low temperatures and $(8-20)\times10^{12}$ ion/cm^2 at room temperature. These data are in good agreement with the experimental results obtained recently.[23] Second, the depths in the surface layer of semiconducting crystals at which track formation is possible can also easily be found. If we base our calculations on the mechanism of thermal peak formation due to strong electron excitation, as a result of which local cylindrical domains melt[24] and are then rapidly quenched, the critical ion energy losses per unit range for track formation in the InP crystal are approximately 1.7–1.9 keV/nm (the upper and lower values of $(dE/dx)_{\mathrm{thr}}$ are given for the case where the process is described with and without allowance for the latent heat of fusion). The information drawn from Fig. 3a (curve *2*) suggests that in the case being discussed, tracks can be formed starting at the crystal's surface to 10–15 μm deep. The experiment of Gaiduk *et al.*[23] showed that there are continuous track regions at depths up to 10–12 μm. The broken tracks at $x\geqslant10$ μm observed in the experiment may be due to statistical fluctuations in the charge-exchange processes involving the loss of one or several electrons by the ion, when in a certain segment of the range dE/dx becomes larger than the value $(dE/dx)_{\mathrm{thr}}$ for track formation. Hence when A^{3}B^5 semiconductors are irradiated by high-energy ions, an extremely complicated pattern of structural transformations over the depth of the crystal can be observed, a pattern that contains continuous track regions, broken tracks, fairly weakly damaged regions, and amorphous regions at the ends of ion ranges in the crystal.

5. CONCLUSIONS

Numerical modeling suggests that fluctuations in the charge states of atoms have a strong effect on the formation of implantation profiles in high-energy ion implantation, especially on the variance and asymmetry of the ion distribution over the depth of the crystal. As the atomic number of the ions increases, the relative size of this effect diminishes. The initial charge state of the ion must be taken into account in the numerical calculations, but its effect is not very strong.

Numerical calculations of the distributions of the energy released by ions in inelastic and elastic processes make it possible to describe such important phenomena as the formation of continuous and broken tracks and latent amorphous layers inside the crystal. Satisfactory agreement is observed between the theoretical predictions and the experimental data on the depths of occurrence and the sizes of these regions in a crystal.

This work was supported by the Foundation for Basic Research of the Republic of Belarus, and the International Soros Science Education Program (ISSEP).

[1] A. La Ferla, A. Difranko, E. Rimini *et al.*, Mater. Sci. Eng. B **2**, 69 (1989).
[2] A. La Ferla, L. Torrisi, G. Galvagno *et al.*, Nucl. Instrum. Methods Phys. Res. B **73**, 9 (1993).
[3] A. F. Burenkov, V. S. Varychenko, A. M. Zaitsev *et al.*, Phys. Status Solidi A **115**, 427 (1989).
[4] A. F. Burenkov, F. F. Komarov, G. G. Konoplyanik *et al.*, Nucl. Instrum. Methods Phys. Res. B **48**, 439 (1990).
[5] A. M. Zaitsev, S. A. Fedorov, A. A. Mel'nikov *et al.*, Nucl. Instrum. Methods Phys. Res. B **82**, 421 (1993).
[6] A. F. Burenkov, F. F. Komarov, S. A. Fedorov *et al.*, Radiat. Eff. **25**, 169 (1993).
[7] A. F. Burenkov, F. F. Komarov, and S. A. Fedorov, Phys. Status Solidi B **169**, 93 (1992).
[8] A. F. Burenkov, F. F. Komarov, and S. A. Fedorov, Nucl. Instrum. Methods Phys. Res. B **67**, 30 (1992).
[9] A. F. Burenkov, F. F. Komarov, and S. A. Fedorov, Nucl. Instrum. Methods Phys. Res. B **67**, 35 (1992).

[10] N. Bohr, Kgl. Danske Videnskab. Selskab. Mat.-Fys. Medd. **18**, No. 8 (1948).
[11] P. Sigmund, Nucl. Instrum. Methods Phys. Res. B **69**, 113 (1992).
[12] A. Narmann and P. Sigmund, Phys. Rev. A **49**, 4709 (1994).
[13] A. Arnau, Nucl. Instrum. Methods Phys. Res. B **93**, 195 (1994).
[14] H. Ogawa, I. Katayama, I. Sugai *et al.*, Phys. Lett. A **167**, 487 (1992); **160**, 77 (1991).
[15] G. G. Bentim, M. Biancom, and R. Nipoti, Nucl. Instrum. Methods Phys. Res. B **80/81**, 33 (1993).
[16] J. E. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Range of Ions in Solids*, Pergamon Press, New York (1985), p. 321.
[17] A. F. Burenkov, F. F. Komarov, M. A. Kumakhov *et al.*, A. F. Burenkov, F. F. Komarov, M. A. Kumakhov *et al.*, *Tables of Ion Implantation Spatial Distribution*, Gordon and Breach, New York (1986), p. 412.
[18] W. Brandt and M. Kitagawa, Phys. Rev. B **25**, 5631 (1982).
[19] H.-D. Betz, Rev. Mod. Phys. **44**, 465 (1972).
[20] F. F. Komarov, A. P. Novikov, A. F. Burenkov, *Ion Implantation* [in Russian], Belorussian University Press, Minsk (1994).
[21] W. Wesch, E. Wendler, and G. Götz, J. Appl. Phys. **65**, 519 (1989).
[22] W. Wesch, Nucl. Instrum. Methods Phys. Res. B **68**, 340 (1992).
[23] P. I. Gaiduk, F. F. Komarov, W. Wesch *et al.*, Phys. Rev. Lett. (in press).
[24] F. F. Komarov, Langmuir **12**, 199 (1996).

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Influence of the Stark effect on multiphoton ionization of atoms when the dynamic polarizability depends strongly on the laser frequency

I. I. Bondar' and V. V. Suran

Uzhgorod State University, 294000 Uzhgorod, Ukraine

(Submitted 17 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 499–512 (February 1998)

The influence of the Stark effect on multiphoton ionization of Ba atoms under conditions when the dynamic polarizability depends strongly on the frequency of the laser radiation is investigated. It is found that for some electric field strengths ε of the laser radiation this effect gives rise to resonance peaks in the Ba^+ ion yield as a function of the laser radiation frequency at frequencies corresponding to single-photon transitions between the excited states. These frequencies can differ substantially from the frequencies corresponding to the conventional multiphoton excitation of these states from the ground states of the atoms. Peaks in the ion yield as a function of ε behave differently from the conventional Stark effect—their position on the frequency scale does not depend strongly on ε. The conditions under which such an induced resonance structure appears are determined. © *1998 American Institute of Physics.* [S1063-7761(98)00802-6]

1. INTRODUCTION

When interpreting the results of investigations of multiphoton ionization of atoms, the perturbation of the atomic spectrum by the electric field of laser radiation must be taken into account in most cases.[1] One of the basic mechanisms by which a strong laser field perturbs an atomic spectrum is the dynamic Stark effect. This effect consists of a change ΔE_n in the energy of a level n (Ref. 1),

$$\Delta E_n = -\alpha_n(\omega)\varepsilon^2/4. \tag{1}$$

Here ε is the intensity of the laser field and $\alpha_n(\omega)$ is the dynamic polarizability of the level n, which depends on the frequency of the radiation,[2]

$$\alpha_n(\omega) = \sum_m \frac{d_{nm}^2}{\omega_{nm} - \omega - i\Gamma_m/2}. \tag{2}$$

In Eq. (2) d_{nm} is the matrix element of a one-photon transition from the state n into the state m; ω_{nm} is a frequency corresponding to this transition; ω is the frequency of the laser radiation; and Γ_m is the width of the level m. The summation in Eq. (2) extends over all intermediate states m.

The character of the manifestation of the Stark effect accompanying multiphoton ionization when the laser radiation frequency differs strongly from the frequencies ω_{nm} corresponding to single-photon transitions in atomic spectra is now well known. In these cases, according to Eq. (2), the magnitude of the dynamic polarizability is virtually independent of the frequency of the laser radiation. For this reason, in the case of multiphoton ionization of atoms under these conditions, according to Eq. (1) the shifts of the resonance peaks in the ion yield $N^+(\omega)$ as a function of the laser frequency should depend quadratically on ε. This picture has indeed been observed in most experiments on the multiphoton ionization of atoms.[1]

If we have $\omega \simeq \omega_{nm}$, then the dynamic polarizability of the levels n and m is strongly frequency-dependent. For this reason, the Stark effect in the case of multiphoton ionization of atoms under such conditions should be different from the picture described above. However, this case has never been investigated in detail.

In the present paper we report the results of investigations of the manifestation of the Stark effect in the ionization of a Ba atom when $\omega \simeq \omega_{nm}$. Radiation from a color-center laser frequency-tunable in the range $\omega = 8400-9000$ cm^{-1} was used in these investigations. The line width of this radiation was equal to $\Delta\omega = 3-4$ cm^{-1} and the duration of the laser pulse was $\tau \simeq 3\cdot 10^{-8}$ s. Ionization of a Ba atom under these conditions requires the absorption of five photons. The condition $\omega = \omega_{nm}$ is satisfied for a number of frequencies of the laser radiation employed.

In the experiment the yield of Ba^+ ions produced through multiphoton ionization of Ba atoms was measured as a function of the laser frequency ω with different intensities ε of the electric field of laser radiation. The intensity of the laser electric field varied in the range $\varepsilon \simeq 10^6 - 7\cdot 10^6$ V/cm. On the whole, the experiments were arranged in the conventional manner for investigating the multiphoton ionization of atoms. The experimental arrangement is described in detail in previous treatments (see, for example, Ref. 3).

2. EXPERIMENTAL RESULTS

The results of the investigations of the dependence of the yield of Ba^+ ions on the frequency of the laser radiation are presented in Figure 1 for three values of ε. Here and below, ε denotes the maximum field strength within a laser pulse. As one can see from Fig. 1, the character of the frequency dependence $N^+(\omega)$ varies with ε. The number of resonance peaks in these dependences does not remain constant. The position of the resonance peaks on the frequency scale like-

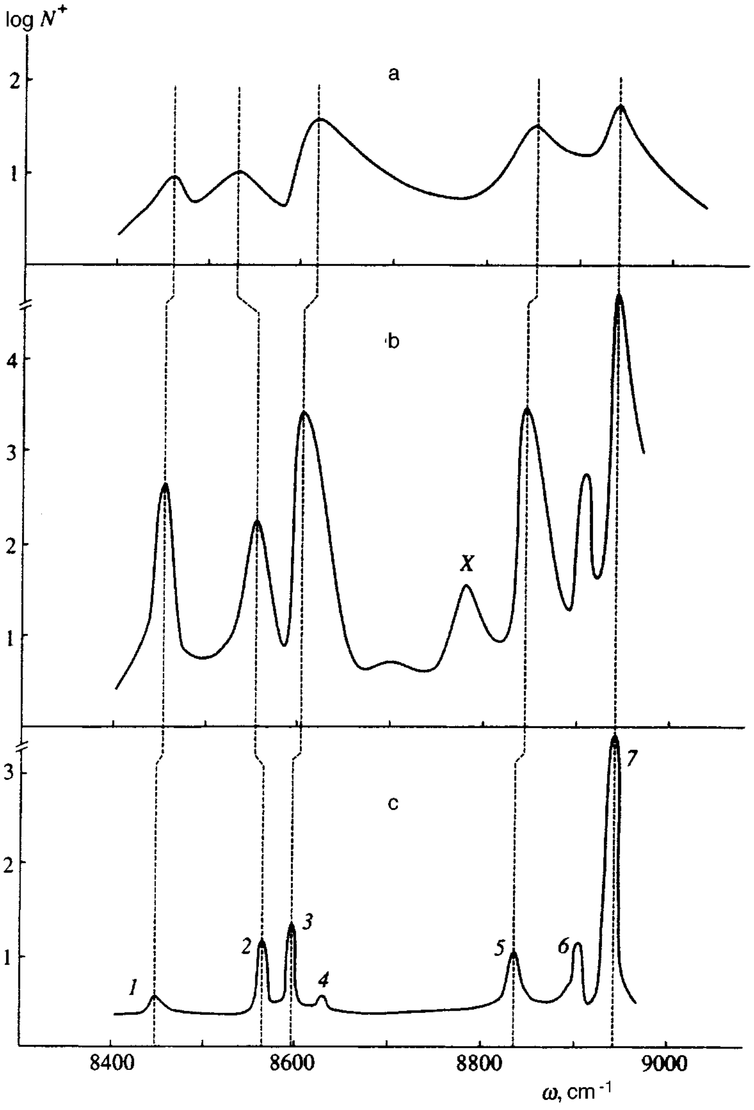


FIG. 1. Yield of Ba^+ ions versus frequency ω of the radiation from a color-center laser. The curves were measured with the strengths ε of the electric field of the laser radiation equal to $6.6\cdot 10^6$ V/cm (a), $4\cdot 10^6$ V/cm (b), and $9\cdot 10^5$ V/cm (c).

wise, in the main, does not remain the same. For large values of ε these peaks show a distinct asymmetry. These facts together indicate that the spectrum of the Ba atom is strongly perturbed under the conditions of our experiments.

We note that the frequency dependence $N^+(\omega)$ measured with approximately the same values of ε as those displayed in Figs. 1a and 1c was analyzed in detail in Refs. 4 and 5. For this reason, we shall not discuss here the details of these dependences. We note only that the results of Refs. 4 and 5 show that the dependence presented in Fig. 1c can be interpreted as the spectrum of unperturbed states of the Ba atom. Table I gives the identification of the resonance peaks present in this dependence. As follows from the table, a series of ordinary multiphoton resonance transitions from the $6s^2\,{}^1S_0$ ground state of the Ba atom as well as one transition, consisting of two successive resonance transitions (the transition at the frequency $\omega = 8940$ cm^{-1}) occur in the experimental region of the spectrum. As the results of our previous work showed,[5] the behavior of the peaks, which are due to conventional multiphoton transitions, with increasing ε is described well by the conventional quadratic Stark effect with $\omega \neq \omega_{nm}$: their shift and width depend quadratically on ε.

We now examine the frequency dependence $N^+(\omega)$ presented in Fig. 1b. It was measured for a value of ε falling between the values of ε at which the curves shown in Figs. 1a and 1c were measured. As one can see, a resonance maximum is present in this dependence at the frequency

TABLE I. Identification of resonance peaks in the frequency dependence $N^+(\omega)$ presented in Fig. 1c.

N	ω, cm^{-1}	ω_1, cm^{-1}	Transition
1	8452	8448	$6s^2\,{}^1S_0 + 4\hbar\omega \rightarrow 5d7s\,{}^1D_2$
2	8565	8568	$6s^2\,{}^1S_0 + 3\hbar\omega \rightarrow 5d6p\,{}^3P_1^0$
3	8595	8592	$6s^2\,{}^1S_0 + 4\hbar\omega \rightarrow 6p^2\,{}^1S_0$
4	8625	8623	$6s^2\,{}^1S_0 + 4\hbar\omega \rightarrow 6p^2\,{}^3P_0$
5	8838	8836	$6s^2\,{}^1S_0 + 4\hbar\omega \rightarrow 6p^2\,{}^1D_2^0$
6	8905	8904	$6s^2\,{}^1S_0 + 4\hbar\omega \rightarrow 6p^2\,{}^3P_2$
7	8940	8939	$6s^2\,{}^1S_0 + 3\hbar\omega \rightarrow 5d6p;\ {}^1F_3^0 + \hbar\omega \rightarrow 6s7d\,{}^3D_2$

Note. The numbers in the table correspond to the numbers of the resonance peaks in the frequency dependence $N^+(\omega)$ in Fig. 1c; ω—laser frequencies at which resonance peaks appear in the curves $N^+(\omega)$; ω_1—frequencies corresponding to a multiphoton transition into undisturbed states of the Ba atom.

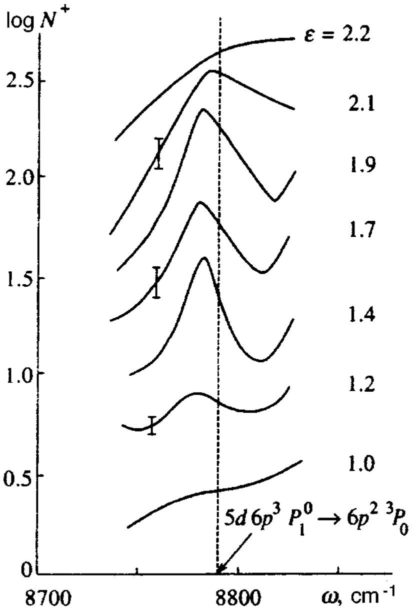


FIG. 2. Results of investigations of the behavior of an induced peak due to the one-photon transition $5d6p\ ^3P_1^0 \rightarrow 6s^2\,^3P_0$ with different field intensities ε. The values of ε are presented in scaled units. The value $\varepsilon = 1$ corresponds to the absolute value $\varepsilon = 2.5 \cdot 10^6$ V/cm. The vertical dashed line marks the frequency corresponding to this one-photon transition.

$\omega = 8780\ \text{cm}^{-1}$; this maximum is not present in the frequency dependence $N^+(\omega)$ in Figs. 1a and 1c (the peak X in Fig. 1b).

To study the behavior of this peak as a function of ε the frequency dependence $N^+(\omega)$ was investigated in detail near the frequency $\omega = 8780\ \text{cm}^{-1}$ by varying ε with a small step. The results of these investigations are presented in Fig. 2. As follows from this figure, the peak noted above first appears at some field intensity ($\varepsilon_0 = 3 \cdot 10^6$ V/cm). The width of this peak is comparatively large relative to the widths of the peaks which are due to the conventional multiphoton excitation of the unperturbed states (see Fig. 1c). As ε varies, the position of this peak on the frequency scale does not change much within the width of the peak. This peak is not observed for large values of ε.

Analysis shows that the peak noted above cannot be identified with a conventional multiphoton excitation of some bound and autoionization states of a Ba atom from either the $6s^2\,^1S_0$ ground state or the first excited states. At the same time, the frequency $\omega = 8780\ \text{cm}^{-1}$ at which this peak appears is close to the frequency corresponding to the one-photon transition $5d6p\ ^3P_1^0 \rightarrow 6p^2\,^3P_0$ ($\omega_{mn} = 8788\ \text{cm}^{-1}$). However, this frequency does not correspond to multiphoton excitation of either the $5d6p\ ^3P_1^0$ state or the $6p^2\,^3P_0$ state. The detuning for multiphoton excitation of these levels at frequency $\omega_{mn} = 8788\ \text{cm}^{-1}$ equals $\Delta E = 664\ \text{cm}^{-1}$ (see Fig. 3a).

At present, two possible causes of the appearance of the resonance structure in the excitation scheme shown in Fig. 3 are known—resonant mixing, leading to splitting of the levels, and the Stark effect, leading to a shift of the levels.[1]

In the first case the detuning ΔE must be compensated by splitting of the states into quasienergy levels:

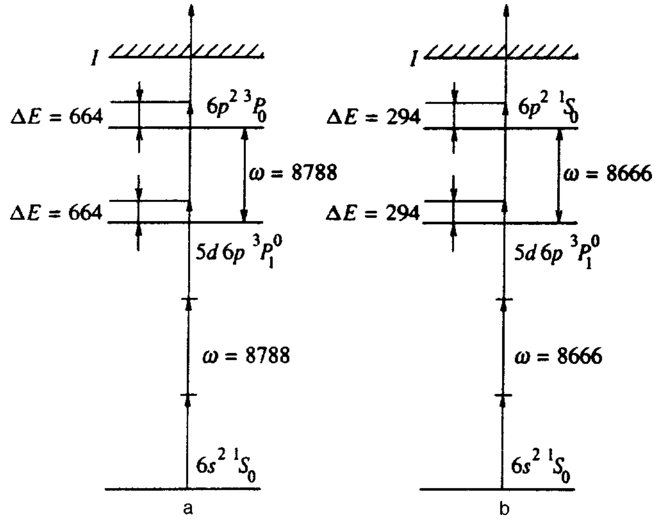


FIG. 3. Diagrams illustrating the realization of the processes investigated. The magnitude ΔE of the detuning is presented for an unperturbed Ba atom. The values of ΔE and ω are given in cm^{-1}.

$$\Delta E = \Omega = d_{nm}\varepsilon/2, \tag{3}$$

where Ω is the Rabi frequency and d_{nm} is the matrix element of a single-photon transition (in our case the transition $5d6p\ ^3P_1^0 \rightarrow 6p^2\,^3P_0$).

However, analysis shows that under our experimental conditions this effect cannot produce a peak in the Ba$^+$ ion yield. The estimates show that for the field strength $\varepsilon = 3 \cdot 10^6$ V/cm given above the splitting can compensate the detuning $\Delta E = 664\ \text{cm}^{-1}$ only if $d_{nm} \geqslant 10$ a.u. This is a very large value of the transition matrix element. It is larger than for the transition from the ground state of the Ba atom into the first excited state $6s6p\ ^1P_1^0$, for which $d_{nm} \simeq 8.6$ a.u.[6] In addition, it is well known that d_{nm} should be much smaller for transitions between excited states, which the transition $5d6p\ ^3P_1^0 \rightarrow 6p^2\,^3P_0$ under discussion is.

Moreover, when resonant mixing occurs the corresponding peak should be symmetric with respect to the resonance frequency $\omega_{mn} = 8788\ \text{cm}^{-1}$. In our case, however, it is detuned from this frequency by the amount $\delta = 8\ \text{cm}^{-1}$. This also contradicts the fact that the peak under discussion is due to resonant mixing of the levels.

Let us now examine the possibility for the Stark effect to balance the detuning. In this case the detuning must be balanced by a level shift

$$\Delta E = (\alpha_n - \alpha_0)\varepsilon^2/4, \tag{4}$$

where α_0 and α_n are the dynamic polarizabilities of the ground and excited states, respectively.

By analogy, we shall estimate the value of d_{nm} required for the Stark effect to balance the detuning. We shall use Eq. (4). This formula contains the shift of the ground state of the atom: $\Delta E_0 = -\alpha_0\varepsilon^2/4$. Estimates show that under our experimental conditions this quantity does not exceed $5\ \text{cm}^{-1}$, which is much less than ΔE. Therefore we shall neglect ΔE_0 in these estimates. We note that in estimating ΔE_0 we took the value of α_0 to be equal to its static limit $\alpha_0 = 280$ a.u.[6] Such an approach is completely justified, since the laser frequencies employed are much lower (approximately by a factor of 2) than the frequencies corresponding to the transition into the first excited state $6s6p\ ^1P_1^0$ ($\omega=18060$ cm^{-1}), near which α_0 should be substantially different from its static limit.

As far as the dynamic polarizability of an excited state is concerned, according to Eq. (2), near frequencies ω_{nm} corresponding to single-photon transitions between two states the dynamic polarizability of the two states involved has poles and can assume large values. The quantity Γ_m ensures that the dynamic polarizability remains finite.

For estimates, we retain in the sum over the intermediate states in the expression (2) for the dynamic polarizability α_n only the resonance term that makes the largest contribution to $\alpha_n(\omega)$. After simple manipulations, we obtain the following expression for the real part of the dynamic polarizability giving rise to a level shift:

$$\alpha_n(\omega)=\frac{(\omega_{nm}-\omega)d_{nm}^2}{(\omega_{nm}-\omega)^2+\Gamma_m^2/4}. \tag{5}$$

We note that in the general case Γ_m in the formulas for the dynamic polarizabilities is taken to be the largest of the natural width of a level, the field and ionization widths of the level, and the width of the radiation spectrum. For definiteness, we take for estimates $\omega_{nm}-\omega\simeq\Gamma_m/2\simeq\delta$, where $\delta=8$ cm^{-1}, as noted above, is the detuning of the peak under discussion relative to the resonance frequency $\omega_{nm}=8788$ cm^{-1}. Then Eq. (5) becomes

$$\alpha_n=d_{nm}^2/2\delta. \tag{6}$$

Thus, on this basis, the equation for estimating the value of δ_{nm} necessary in order for the Stark effect to compensate the detuning ΔE is

$$d_{nm}=\sqrt{8\delta\Delta E/\varepsilon^2}. \tag{7}$$

An estimate based on this expression for the conditions of our experiment gives $d_{nm}\simeq1$ a.u. This value, in contrast to the value obtained above for the case of resonance mixing, is typical for transitions between excited states. This confirms that under the conditions of the experiments described above the Stark effect could balance the detuning ΔE and thereby give rise to the peak being discussed.

We note that all aspects of this peak are explained well by the occurrence of a Stark effect under conditions when the dynamic polarizability depends strongly on the frequency of the laser radiation. Thus, the asymmetry of this peak relative to the resonance frequency $\omega_{nm}=8788$ cm^{-1} is explained well. We shall explain this fact using Fig. 4. It is well known that for a two-level system in a field with frequencies close to the resonance frequency ω_{nm} for $\omega\geqslant\omega_{nm}$ one level possesses a negative and the other a positive dynamic polarizability; for $\omega\leqslant\omega_{nm}$ the polarizabilities of both levels change sign[1] (see Fig. 4a). Thus, one level has a positive shift for $\omega\geqslant\omega_{nm}$ and the other has a positive shift for $\omega\leqslant\omega_{nm}$ (see Fig. 4b). Therefore, for $\omega\geqslant\omega_{nm}$ the first level will be excited because the detuning ΔE is compensated, while for $\omega\leqslant\omega_{nm}$ the second level will be excited.

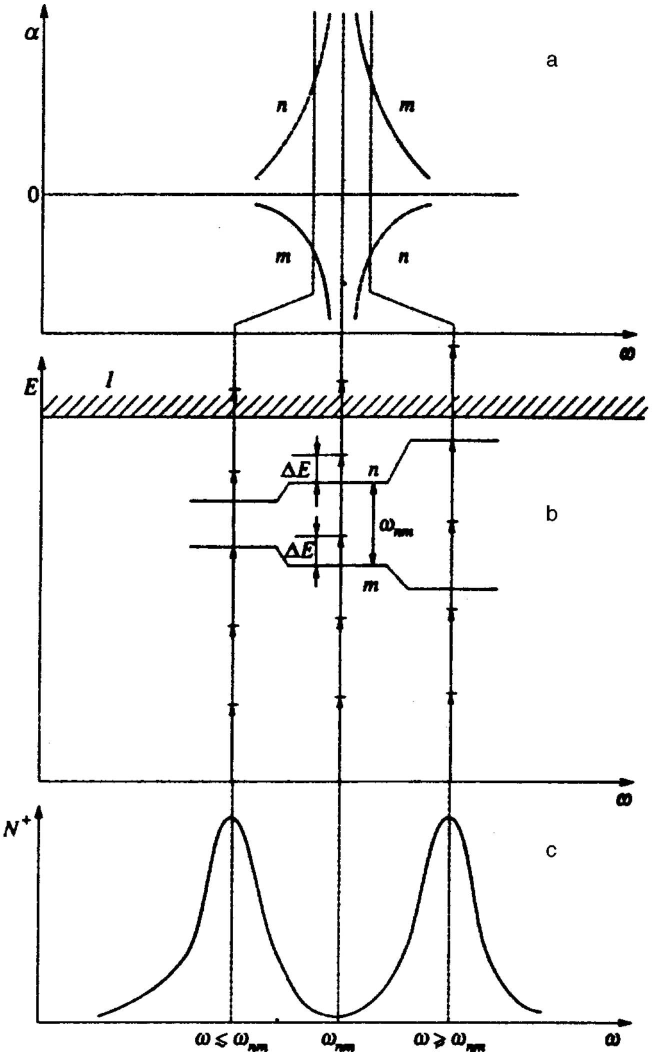


FIG. 4. Diagram of the formation of the Stark-effect-induced resonance structure in $N^+(\omega)$ that corresponds to the excitation schemes presented in Fig. 3: a—Approximate dependence of the dynamic polarizability α of the excited n and m states on the laser frequency near the frequency ω_{nm} corresponding to a single-photon transition between these states; b—diagram of the change in the energy E of the states n and m and appearance of induced resonances with these states; c—approximate form of the resonance structure in the frequency dependence $N^+(\omega)$ in the case when the multiphonon excitation probabilities of the states n and m are the same.

When exact equality $\omega=\omega_{nm}$ holds the dynamic polarizabilities of both levels equal zero. In this case neither level is shifted and therefore neither level is perturbed.

For this reason, when the probabilities of multiphoton excitation of both interacting levels are the same, two peaks with the same amplitude and a dip between them at $\omega=\omega_{nm}$ should appear near the frequency ω_{nm} (see Fig. 4c). However, if the excitation probabilities of both levels are substantially different, then, naturally, these two peaks will have very different amplitudes.

As follows from Fig. 1c, the resonance peak due to three-photon excitation of the $5d6p\ ^3P_1^0$ state ($\omega=8565$ cm^{-1}) has a much larger amplitude than the peak due to four-photon excitation of the state $6p^{2\,3}P_0$ ($\omega=8525$ cm^{-1}). This shows that the probability of three-photon excitation of the $5d6p\ ^3P_1^0$ state is much higher than that of four-photon excitation of the $6p^{2\,3}P_0$ state. Hence it follows that the peak at $\omega = 8780\ \mathrm{cm}^{-1}$ is due to three-photon excitation of the $5d6p\ ^3P_1^0$ state strongly shifted by the Stark effect.

The first appearance of this peak at some value of ε_0 is explained by the fact that the detuning of the multiphoton resonance is first balanced at this field intensity, i.e., the following condition is now satisfied:

$$\Delta E = -(\alpha_{\max} - \alpha_0) F_0^2/4, \tag{8}$$

where $\alpha_{\max}$ is the maximum magnitude of the dynamic polarizability of the shifted level near the resonance frequency ω_{nm} and F is the effective field strength for multiphoton excitation of the atom, i.e., the field strength corresponding to the part of the laser–atom volume that makes the largest contribution to the total number of excited atoms. Since the distribution of the laser radiation is nonuniform, we have $F = f\varepsilon$, where as noted above ε is the maximum intensity of the electric field of the laser radiation within the laser pulse and $f < 1$ is a coefficient. In Eq. (8) $F_0 = f\varepsilon_0$.

Calculations taking account of the real space–time distribution of the laser radiation in our experiments show that we have $f = 0.80$ for three-photon excitation.

The maximum value of the dynamic polarizability of the $5d6p\ ^3P_1^0$ level near the frequency $\omega = 8780\ \mathrm{cm}^{-1}$ was calculated according to Eq. (8). The result is $\alpha_{\max} = -3.4 \cdot 10^4$ a.u.

The dynamic polarizability of the same level $5d6p\ ^3P_1^0$ was measured in Ref. 5 for the frequency corresponding to three-photon excitation ($\omega_0 = 8568\ \mathrm{cm}^{-1}$). It equals $\alpha = 2.26 \cdot 10^3$ a.u. Comparing these two values of the dynamic polarizability of the $5d6p\ ^3P_1^0$ level shows that for a comparatively small change in the frequency of the laser radiation the dynamic polarizability changes sign and by more than an order of magnitude. This fact shows that in the spectral range studied the dynamic polarizability depends strongly on the frequency of the laser radiation.

As noted in the introduction, the Stark effect in multiphoton ionization conventionally manifests itself as a quadratic increase with ε of the detuning of the resonance peaks in $N^+(\omega)$ relative to their unperturbed positions. In the case at hand, however, as ε increases, the position of the resonance peak under discussion on the frequency scale virtually does not change within the width of the peak. This effect is explained by the sharp dependence of the polarizability on the frequency of the laser radiation near ω_{nm}. We shall explain this in greater detail.

As noted above, the induced resonance peak first appears when the relation (8) is satisfied. As the strength of the field increases further, because of the fact that when $N^+(\omega)$ is measured the magnitude of the dynamic polarizability changes together with the laser frequency, the resonance maximum should appear at a frequency where $|\alpha| < |\alpha_{\max}|$ holds and which satisfies the relation

$$\omega = \omega_0 + \Delta E'/k = \omega_0 - \alpha_n(\omega) f^2 \varepsilon^2 / 4k, \tag{9}$$

where k is the number of photons required for excitation of the level and ω_0 is the frequency corresponding to this excitation. Specifically, this happens at $k = 3$ and $\omega_0 = 8568\ \mathrm{cm}^{-1}$ for the $5d6p\ ^3P_1^0$ level. In Eq. (9) the perturbation of the ground state of the Ba atom was once again neglected.

We note that in the case of excitation of the state $5d6p\ ^3P_1^0$ undergoing a shift, we have $\Delta E' < \Delta E$ because of the sign of the dynamic polarizability (as noted above, ΔE is the detuning at which the induced resonance peak first appears), i.e., in contrast to the conventional manifestation of the Stark effect, in the present case the detuning should decrease and not increase with increasing ε. However, on account of the sharp frequency dependence of $\alpha_n(\omega)$ the value of $\Delta E'$ will not differ much from ΔE as ε varies over a large range.

To confirm this, we shall solve Eq. (9) for ω. For this, we replace $\alpha_n(\omega)$ in this equation by the real part of the dynamic polarizability, consisting only of the resonance term (5). Once again, we set $\omega - \omega_{nm} \simeq \Gamma/2$ for frequencies near the frequency corresponding to a single-photon transition. After manipulations we obtain the following frequency dependence of the dynamic polarizability:

$$\alpha_n(\omega) = \frac{C}{2(\omega - \omega_{nm})}, \tag{10}$$

where C is a constant. We determine its value from the value of the dynamic polarizability $\alpha_{\max}$ at the frequency $\omega = 8790\ \mathrm{cm}^{-1}$. The result is $C = -2.7 \cdot 10^{-8}$ a.u.

Using the expression obtained above for $\alpha_n(\omega)$, we find that the equation (9) for the frequency at which a maximum should as a function of ε becomes

$$(\omega - \omega_0)(\omega - \omega_{nm}) + C'\varepsilon^2 = 0, \tag{11}$$

where $C' = Cf^2/8k$. The solution of this equation for ω, satisfying the condition $\omega - \omega_{nm} \sim \Gamma_m/2$, has the form

$$\omega = \frac{\omega_{nm} + \omega_0}{2} + \sqrt{\frac{(\omega_{nm} - \omega_0)^2}{4} - C'\varepsilon^2}. \tag{12}$$

Estimates show that under our experimental conditions $C'\varepsilon^2 \ll (\omega_{nm} - \omega_0)^2/4$. Hence we find that the frequency ω at which an induced resonance peak should appear is virtually independent of the field strength for $\varepsilon > \varepsilon_0$ and $\omega \simeq \omega_{nm}$.

Thus, although the formulas employed above are substantially simplified, they nonetheless describe the observed effect well.

Note that the change in frequency ω can be observed in principle, but only for a large change in ε. Thus, in accordance with Eq. (12), in order for the frequency ω to change relative to ω_{nm} by an amount of the order of the width of the spectrum ($\simeq 5\ \mathrm{cm}^{-1}$), ε must increase to $\varepsilon = 3.5 \cdot 10^8$ V/cm.

However, saturation of the ionization process prevents the observation of such a change in the frequency ω. It is evident that the frequency ω will change with increasing ε until ε equals the value ε_s corresponding to the onset of ionization saturation. As ε increases further, ω will no longer change, since on account of ionization saturation, when the field intensity reaches values $\varepsilon > \varepsilon_s$ there will no longer be any neutral atoms in the interaction region. Thus, the effec-

TABLE II. Single-photon transitions between excited states of the Ba atom which can be realized in the region of the spectrum 8400–9000 cm^{-1}.

ω_2, cm^{-1}	Transition	E, cm^{-1}
8528	$5d6p\ {}^1F_3^0 \rightarrow 6p^2\ {}^1D_2$	1232
8666	$5d6p\ {}^3P_1^0 \rightarrow 6p^2\ {}^1S_0$	294
8788	$5d6p\ {}^3P_1^0 \rightarrow 6p^2\ {}^3P_0$	664
8846	$5d6p\ {}^3D_3^0 \rightarrow 5d7s\ {}^1D_2$	1588

tive value of F with $\varepsilon > \varepsilon_s$ cannot increase, but instead remains equal to $F_s = f\varepsilon_s$. We note that in our case $\varepsilon_s = 5.5 \cdot 10^6$ V/cm.

Note that that appearance of ionization saturation should also result in suppression of the induced peak above. This is explained by the fact that on account of the spatial nonuniformity of the distribution of the laser radiation in the interaction region, as ε increases ($\varepsilon > \varepsilon_0$), the volume where the effective field $F_s = f\varepsilon_s$ is localized will occupy a smaller and smaller part compared with the total volume where ionization of atoms occurs. For this reason, as ε increases under saturation conditions the amplitude of the induced peak should decrease and the peak should vanish.

Thus, analysis shows that a resonance peak of the nature discussed above should appear for $\varepsilon_0 < \varepsilon < \varepsilon_s$, and on account of the sharp frequency dependence of the dynamic polarizability it should appear only near the frequency ω_{nm}. This picture is observed in our experiment.

We note that besides the single-photon transition $5d6p\ {}^3P_1^0 \rightarrow 6p^2\,{}^3P_0$ discussed above, another series of single-photon transitions of the same character between the excited states of the Ba atom can appear in the experimental region of the spectrum. However, the peaks corresponding to these transitions are absent in the frequency dependence $N^+(\omega)$ presented in Fig. 1. This fact can be explained as follows.

As follows from the analysis presented above, the appearance of such peaks depends on the ratio of different factors: the magnitudes of the detuning ΔE for multiphoton excitation of the levels; the magnitudes of the dynamic polarizabilities of the interacting levels; the multiphoton excitation probabilities of these states; and the values of ε_s at which ionization saturation starts.

It is obvious that as ε increases, the peaks for which the smallest detuning for multiphoton excitation or the largest dynamic polarizability of the interacting levels occurs should appear first. The single-photon transitions that can occur in our experimental region of the spectrum with small detunings ΔE are presented in Table II. According to this table, the detuning for the transition $5d6p\ {}^3P_1^0 \rightarrow 6p^2\,{}^1S_0$ ($\Delta E = 294\ \mathrm{cm}^{-1}$), which should occur at the frequency $\omega_{nm} = 8666\ \mathrm{cm}^{-1}$, is smaller than the detuning for the transition $5d6p\ {}^3P_1^0 \rightarrow 6p^2\,{}^3P_0$ discussed above. The corresponding excitation scheme is presented in Fig. 3b. The absence of a corresponding induced peak in the curve $N^+(\omega)$ presented in Fig. 1 shows that the maximum dynamic polarizabilities of the levels corresponding to this transition are much smaller than the dynamic polarizabilities of the levels corresponding to the transition $5d6p\ {}^3P_1^0 \rightarrow 6p^2\,{}^3P_0$.

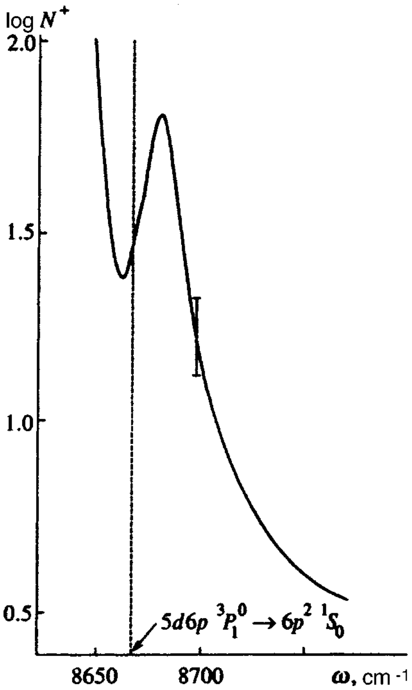


FIG. 5. Yield of Ba^+ ions near the laser frequency corresponding to the single-photon transition $5d6p\ {}^3P_1^0 \rightarrow 6s^2\,{}^1S_0$. The vertical dashed line marks the frequency corresponding to this transition. The measurements were performed for the laser electric field intensity $\varepsilon = 4.5 \cdot 10^5$ V/cm.

If this assumption is correct, then the corresponding peak should appear at values of ε higher than those corresponding to Fig. 1b. For this reason, we searched for such a peak for $\varepsilon > 4 \cdot 10^6$ V/cm. The results are presented in Fig. 5. As follows from this figure, such a peak does appear in the frequency dependence $N^+(\omega)$ under the indicated conditions.

Unfortunately, detailed investigations of the behavior of this peak as a function of ε, similar to those performed for the peak described above, could not be performed because for large ε this peak merges with a neighboring peak.

It is evident from Fig. 5 that, as discussed above, this peak is likewise shifted somewhat relative to the frequency indicated in Table II for the single-photon transition $5d6p\ {}^3P_1^0 \rightarrow 6p^2\,{}^1S_0$. As follows from Fig. 1c, the amplitude of the peak due to four-photon excitation of the $6p^2\,{}^1S_0$ state is larger than the amplitude of the peak due to three-photon excitation of the $5d6p\ {}^3P_1^0$ state. This implies that the probability of four-photon excitation of the $6p^2\,{}^1S_0$ state is higher than that of three-photon excitation of the $5d6p\ {}^3P_1^0$ state. For this reason, the appearance of the peak in question is explained by four-photon excitation of the shifted $6p^2\,{}^1S_0$ state.

The results presented in Fig. 5 make it possible to estimate, as was done for the $5d6p\ {}^3P_1^0$ level, the maximum magnitude of the negative dynamic polarizability of the $6p^2\,{}^1S_0$ level at frequencies near $8666\ \mathrm{cm}^{-1}$. An estimate gives $\alpha_{\max} = -7 \cdot 10^3$ a.u. This value is much less in magnitude than $\alpha_{\max}$ for the $5d6p\ {}^3P_1^0$ level. This is because in the present case the transition whose matrix element makes the largest contribution to α is a triplet–singlet transition, while in the case discussed above it is a triplet–triplet transition. Naturally, the first matrix element is much smaller than the second one.

The dynamic polarizability of the same $6p^{2\,1}S_0$ state was measured in Ref. 5 at a frequency corresponding to four-photon excitation of the state ($\omega_0 = 8592$ cm^{-1}). The results of Ref. 5 imply $\alpha = -3.4 \cdot 10^{-3}$ a.u. As one can see, for the $6p^{2\,1}S_0$ state a small change in the laser frequency also produces a sharp change in the value of α.

As far as the other single-photon transitions presented in Table II are concerned, according to what we have said above the corresponding peaks in the frequency dependence $N^+(\omega)$ should appear at higher values of ε. However, detailed investigations could not be performed in this case, since for large values of ε other broadened peaks, which are due to the conventional multiphoton excitation of the states, are localized at the corresponding frequencies. At the same time, we recall that, as was shown above, ionization saturation strongly affects the appearance of peaks of this kind, which correspond to single-photon transitions for which ΔE is large. Such peaks should not appear in cases when $\varepsilon > \varepsilon_s$.

Let us now consider the frequency dependence $N^+(\omega)$ presented in Fig. 1b. As follows from the analysis presented above, in this dependence two resonance peaks correspond to three-photon excitation of the same $5d6p\ ^3P_1^0$ state. One of these peaks lies near the frequency corresponding to the conventional three-photon excitation of this state ($\omega_0 = 8568$ cm^{-1}). As shown in Ref. 5, the behavior of this peak with increasing ε is conventional for the quadratic Stark effect under conditions of multiphoton ionization: The shift and width are quadratic functions of ε.

The second peak lies near the frequency corresponding to a single-photon transition from the $5d6p\ ^3P_1^0$ state into the higher-energy $6p^{2\,3}P_0$ state ($\omega_{nm} = 8788$ cm^{-1}). As one can see, this frequency differs substantially (by 664 cm^{-1}) from the frequency corresponding to the conventional three-photon excitation of the $5d6p\ ^3P_1^0$ state. The appearance of this peak can be explained by the excitation of the strongly Stark-shifted $5d6p\ ^3P_1^0$ state. The strong shift in this case is due to the large dynamic polarizability of this state. Its behavior with increasing ε (in contrast to the quadratic ε dependence of the shift of the peak) is explained by the sharp dependence of the polarizability on the frequency of the laser radiation.

A similar pattern also obtains for the above-discussed four-photon excitation of the $6p^{2\,1}S_0$ state. In this case, two resonance peaks are likewise present in the frequency dependence $N^+(\omega)$: one peak near the frequency corresponding to conventional four-photon excitation of the state ($\omega_0 = 8592$ cm^{-1}) and the other, Stark-induced, peak near the frequency corresponding to a one-photon transition from this level to the lower-lying $5p6p\ ^3P_1^0$ level ($\omega_{nm} = 8692$ cm^{-1}).

We note that in the general case more than two induced peaks can correspond to the same state. However, their appearance is determined by the relations discussed above between the following quantities characterizing the atom being investigated and the laser radiation: the magnitude of the detuning ΔE for multiphoton excitation of this state; the magnitude of the dynamic polarizability and the probability of multiphoton excitation of this state; and the value of ε_s at which ionization saturation first appears.

3. CONCLUSIONS

In the present work the influence of the Stark effect on multiphoton ionization of atoms under conditions when the dynamic polarizability is strongly frequency dependent was investigated. It was found that for some values of the electric field intensity this influence results in the appearance of induced resonance peaks in the frequency dependence of the ion yield near frequencies corresponding to single-photon transitions between the excited states. These frequencies can be substantially different from the frequencies corresponding to conventional multiphoton excitation of these states from the ground state of the atoms. The behavior of these peaks differs substantially from that of the typical Stark effect—the dependence of their shift on the field strength is far from quadratic (in the case investigated the shift does not depend strongly on ε). The appearance of such peaks depends on the magnitudes of the dynamic polarizabilities of the levels, the multiphoton excitation probabilities of the levels, and the field strengths at which saturation of the multiphoton ionization process first appears.

When the above-noted excitation of one or another state occurs, two and more resonance peaks can appear in the frequency dependence $N^+(\omega)$. One of them appears near the frequency corresponding to the conventional multiphoton excitation of this state from the ground state and the others appear near the frequencies corresponding to single-photon transitions from this state into other states.

In the present work these results were obtained for five-photon ionization of the Ba atoms. We obtained similar results earlier with six-photon ionization of the Sr atom.[7] In this case, likewise for definite values of ε, a resonance peak appears in the frequency dependence $N^+(\omega)$ at a frequency that corresponds not to multiphoton excitation of unperturbed Sr states from the ground state but rather to a one-photon transition between excited states.

These facts show that induced peaks of the nature described above characteristically appear when an atom interacts with laser radiation whose frequency equals the frequencies corresponding to transitions between excited states.

Note that this behavior of the Stark effect is characteristic of the process of multiphoton ionization of atoms by one laser beam, i.e., the exciting field and probe field (which ionizes the atoms) are one and the same. When these fields are different, however, this effect will arise conventionally: The position of the resonance peaks in the curves of the ion yield versus the frequency of the probe radiation should be a quadratic function of the field intensity of the perturbing radiation.

We wish to expression our sincere appreciation to N. B. Delone for his steadfast interest in these investigations.

[1] N. B. Delone and A. P. Kraĭnov, *Atoms in Strong Light Fields*, Springer-Verlag, N. Y., 1985 [Russian original, Moscow, Atomizdat, 1985].

[2] L. P. Rapoport, B. A. Zon, and N. L. Manakov, *Theory of Multiphoton*

Processes in Atoms [in Russian], Atomizdat, Moscow, 1978.

[3] M. V. Ammosov, N. B. Delone, M. Yu. Ivanov, I. I. Bondar, and A. V. Masalov, Adv. At., Mol., Opt. Phys. **29**, 33 (1992).

[4] V. V. Suran and I. I. Bondar', Opt. Spektrosk. **61**(1), 24 (1986) [Opt. Spectrosc. (USSR) **61**, 15 (1986)].

[5] I. I. Bondar' and V. V. Suran, Opt. Spektrosk. (1997), in press.

[6] A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions*, Springer-Verlag, N. Y., 1985 [Russian original, Atomizdat, Moscow, 1980].

[7] I. I. Bondar' and V. V. Suran, Opt. Spektrosk. **72**(1), 9 (1992) [Opt. Spectrosc. (USSR) **72**, 4 (1992)].

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Self-action of intense electromagnetic radiation in an electron–positron vacuum

N. N. Rozanov

Scientific Research Institute of Laser Physics, S. I. Vavilov State Optical Institute, 199034 St. Petersburg, Russia; St. Petersburg State Institute of Precise Mechanics and Optics, 197101 St. Petersburg, Russia

(Submitted 12 March 1997)
Zh. Éksp. Teor. Fiz. **113**, 513–520 (February 1998)

An analysis is done of the effects of self-action of intense coherent electromagnetic radiation in an electron–positron vacuum that is in homogeneous electric and magnetic fields. A modified version of the Heisenberg–Euler theory, in which the Lagrangian incorporates terms with field derivatives, is used to take into account vacuum dispersion. The nonphysical branch of the solutions of the dispersion equation is excluded by a transition to a quasioptical equation for the slowly varying field envelope, an equation that describes the propagation of radiation with allowance for diffraction, spatial–temporal dispersion, and vacuum nonlinearity. The existence of dark solitons (with an intensity gap) in the vacuum is shown to be present. Finally, self-focusing of radiation in a vacuum is demonstrated and the critical self-focusing power is determined. © *1998 American Institute of Physics.* [S1063-7761(98)00902-0]

1. INTRODUCTION

Vacuum polarization in strong electromagnetic fields destroys the linearity of the Maxwell equations.[1] This is reflected in many nonlinear optics phenomena, which can be observed in astrophysical studies[2–6] and in laboratory experiments involving the use of powerful lasers.[7–9] The Heisenberg–Euler theory of this effect leads to Maxwell equations that coincide in form with the equations of nonlinear electrodynamics of continuous media in the absence of dispersion.[1] In this way the vacuum is interpreted as a transparent medium with a weak nonresonant nonlinearity. More precisely, the dispersion is weak far from a resonance for radiation with characteristic frequencies $\omega \ll \omega_c$, where $\omega_c = mc^2/2\hbar$ (m is the electron mass, and c is the speed of light) The absorption of radiation in the vacuum corresponding to electron–positron pair production is exponentially low in fields whose strength is below the critical value, $E \ll E_c = m^2c^3/e\hbar$, where e is the electron charge.

However, the locally weak nonlinearity and dispersion effects can build up along the path of propagation of the radiation. This justifies the importance of taking such effects into account. To incorporate even the weakest dispersion into the picture, we must step outside the framework of the Heisenberg–Euler theory, which was developed for static homogeneous fields. This can be done by introducing terms that contain field derivatives into the Lagrangian.[10,11] But, as it can easily be shown (see Ref. 12 and the discussion of Eqs. ((6)–(10) below), even with the weakest field the dispersion relation acquires new nonphysical branches. The goal of the present paper is to derive and analyze a quasioptical equation for the field envelope that is free of such difficulties (the slowly varying field approximation). The new equation allows for the effects of weak nonlinearity and dispersion of the vacuum, which makes it possible to study a broad range of nonlinear-optics phenomena, such as self-focusing, solitons, and stabilization of shock waves. A brief discussion of the dispersion spread of pulses of weak radiation in a vacuum can be found in Ref. 12. As far as we know, the literature contains no studies of self-action effects in a dispersive vacuum.

2. GENERAL FORMULAS

Below we will use a system of units in which the speed of light c is equal to unity and so is Planck's constant, $\hbar = 1$, while the square of the electron charge e gives the fine structure constant, $e^2 = 1/137$. We start with the electromagnetic-field Lagrangian of the type

$$L = L_0 + L_{HE} + L'. \tag{1}$$

Here $L_0 = -(1/16\pi)F_{ik}F^{ik}$ is the Lagrangian of the classical electromagnetic field; it is quadratic in the electromagnetic field tensor F_{ik} (Ref. 13; here we use the real form of this tensor). The term L_{HE} represents the polarization correction in the Heisenberg–Euler theory (it contains higher powers of F_{ik}; see Ref. 1). Finally, L' is the polarization correction term (found in Ref. 10), which contains derivatives of the electromagnetic field tensor. In an approximation that is lowest in the powers of the field and derivatives we can write

$$L_{HE}^{(4)} = \frac{e^4}{360\pi^2 m^4}(4F^2 + 7G^2), \tag{2}$$

$$L' = g[-(\partial_i F_k^i)(\partial_n F^{nk}) + F_{ik}\Box F^{ik}], \tag{3}$$

where $g = e^2/360\pi m^2$, $\Box = \partial^2/\partial t^2 - \Delta$ is the d'Alembertian, with Δ the Laplace operator, and F and G stand for the following field invariants:

$$F = \tfrac{1}{2}(\mathbf{B}^2 - \mathbf{E}^2), \quad G = \mathbf{E}\cdot\mathbf{B}, \tag{4}$$

with $\mathbf{E}$ and $\mathbf{B}$ the electric and magnetic fields. Such a Lagrangian can be used if

$$\omega \ll m, \quad k \ll m, \quad E \ll E_c, \tag{5}$$

where ω and k are the characteristic frequency and wave vector of the radiation, and $E_c=m^2/e$ is the critical field strength at which electron–positron pair production is still possible.

The first pair of Maxwell equations gives the relationship between the electric and magnetic fields **E** and **B** and the potentials:

$$\operatorname{div}\mathbf{B}=0,\quad \operatorname{curl}\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t}. \tag{6}$$

The second pair can be found by varying the action with the Lagrangian (1), which means we can write this pair in the common form

$$\operatorname{curl}\mathbf{H}=\frac{\partial\mathbf{D}}{\partial t},\quad \operatorname{div}\mathbf{D}=0, \tag{7}$$

where

$$\mathbf{D}=\mathbf{E}+4\pi\mathbf{P},\quad \mathbf{H}=\mathbf{B}-4\pi\mathbf{M}, \tag{8}$$

with $\mathbf{P}=\partial(L_{HE}+L')/\partial\mathbf{E}$ and $\mathbf{M}=\partial(L_{HE}+L')/\partial\mathbf{H}$ the electric and magnetic polarization vectors. Here, however, these quantities are linked to the field strength nonlocally, in contrast to the Heisenberg–Euler theory, in which there is no dispersion.[1] In the lowest approximation in dispersion and nonlinearity we have

$$\mathbf{P}=\mathbf{P}_{HE}^{(3)}-6g\Box\mathbf{E},\quad \mathbf{H}=\mathbf{H}_{HE}^{(3)}+6g\Box\mathbf{B}, \tag{9}$$

$$\begin{aligned}\mathbf{P}_{HE}^{(3)}&=\frac{e^4}{180\pi^2m^4}(-4F\mathbf{E}+7G\mathbf{B}),\\ \mathbf{M}_{HE}^{(3)}&=\frac{e^4}{180\pi^2m^4}(-4F\mathbf{B}+7G\mathbf{E}).\end{aligned} \tag{10}$$

The Maxwell equations (6) and (7) in conjunction with the constitutive equations (8)–(10) form a closed system of equations describing the nonlinear and dispersion effects in the propagation of electromagnetic radiation in an electron–positron vacuum, provided that conditions (5) are met. Note that in accordance with (9) dispersion is both spatial and temporal and is characterized by a single constant g. In view of the first two inequalities in (5), dispersion must be weak. Ignoring this requirement leads to the following difficulty. When we are dealing with extremely weak fields, $E\ll(\omega/m)E_c$, in the expressions (9) for the polarization vectors we could ignore the terms that are nonlinear in the field strength, $P_{HE}^{(3)}, H_{HE}^{(3)}\rightarrow 0$, and leave only the dispersion terms. But, in accordance with the wave equation, in addition to the ordinary branch $k^2=\omega^2$ there would then appear a new branch of the dispersion with no physical meaning. Corresponding to this branch are rapid spatial–temporal variations of the field that do not meet the conditions (5).

As in the Heisenberg–Euler theory, a plane wave is still a solution of the reduced equations, i.e., it does not polarize the vacuum. Although formally the additional (''dispersive'') terms in (9) are linear in the field, they disappear in the case of weak fields satisfying the ordinary dispersion relation. Vacuum dispersion manifests itself when the classical dispersion relation $k^2=\omega^2$ is violated. This can be achieved by various means: due to the nonlinearity in the interaction of several plane waves (crossed light beams),[7–9] in the presence of a rarefied medium (plasma),[14] etc. Below we consider the simplest possible variant of propagation of electromagnetic radiation against the background of strong static homogeneous electric and magnetic fields that violate the above dispersion relation.

3. WEAK RADIATION AGAINST THE BACKGROUND OF STATIC FIELDS

Suppose that the field is a mixture of static electric and magnetic fields (labeled by the subscript 0) and a beam/pulse of relatively high-frequency electromagnetic radiation (primed quantities):

$$\mathbf{E}=\mathbf{E}_0+\mathbf{E}',\quad \mathbf{B}=\mathbf{B}_0+\mathbf{B}'. \tag{11}$$

In this section we assume that the high-frequency radiation is weak and ignore the terms that are nonlinear in $\mathbf{E}'$ and $\mathbf{B}'$ in the expressions for the polarization vectors:

$$\mathbf{D}=\mathbf{D}_0+\mathbf{D}',\quad \mathbf{H}=\mathbf{H}_0+\mathbf{H}'. \tag{12}$$

Weak radiation can be represented by a superposition of weak plane waves with a spatial–temporal dependence of the form

$$\mathbf{E}'=\mathbf{E}_1\exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)], \tag{13}$$

with similar expressions for $\mathbf{B}'$, $\mathbf{D}'$, and $\mathbf{H}'$. For an individual plane wave characterized by a wave vector **k** and frequency ω the Maxwell equations reduce to

$$\mathbf{k}\times\mathbf{E}'=\omega\mathbf{B}',\quad \mathbf{k}\times\mathbf{H}'=-\omega\mathbf{D}'. \tag{14}$$

The constitutive equations, which express $\mathbf{D}'$ and $\mathbf{H}'$ in terms of $\mathbf{E}'$ and $\mathbf{B}'$, can be obtained by linearizing the expressions (9) and (10) in $\mathbf{E}'$ and $\mathbf{B}'$. These equations readily yield dispersion relations between the wave vector **k** and frequency ω. External static fields lead to vacuum anisotropy, which make the general relationships cumbersome. To simplify matters, we consider the case common in the optics of uniaxial crystals: the presence of only a static magnetic field $(\mathbf{E}_0=0)$. Then the constitutive equations become

$$\begin{aligned}\mathbf{D}'&=(1-\beta)\mathbf{E}'+7\gamma(\mathbf{E}'\cdot\mathbf{B}_0)\mathbf{B}_0,\\ \mathbf{H}'&=(1-\beta)\mathbf{B}'-4\gamma(\mathbf{B}'\cdot\mathbf{B}_0)\mathbf{B}_0,\end{aligned} \tag{15}$$

where

$$\beta=2\gamma+24\pi g(k^2-\omega^2),\quad \gamma=\frac{e^4}{45\pi m^4}. \tag{16}$$

This corresponds to tensors of electric and magnetic permittivity of the vacuum, i.e.,

$$D_i'=\epsilon_{ik}E_k',\quad H_i'=\mu_{ik}B_k', \tag{17}$$

of the form

$$\epsilon_{ik}=1-\beta+7\gamma B_{0i}B_{0k},\quad \mu_{ik}=1-\beta-4\gamma B_{0i}B_{0k}. \tag{18}$$

Spatial–temporal dispersion manifests itself in the dependence of β on the frequency ω and the wave number k. If this dependence is ignored, the result coincides with the one in Ref. 1 (provided that the error in Eq. (130.4) in Ref. 1 is corrected). Accordingly, there is no vacuum anisotropy and dispersion in an external field if $\mathbf{B}'\cdot\mathbf{B}_0=0$ and at the same time either $\mathbf{E}'\cdot\mathbf{B}_0=0$ or $\mathbf{k}\times\mathbf{B}_0=0$. In all other cases there is vacuum anisotropy and dispersion. A convenient way to describe these phenomena is to select for the two independent polarization the linear polarizations in which the vector $\mathbf{B}'$ is perpendicular to the plane containing $\mathbf{k}$ and $\mathbf{B}_0$ (labeled $\perp$) or in which it lies in that plane (labeled $\|$). Introducing a small deviation of the refractive index from unity, δn, calculated at a fixed polarization in the absence of dispersion, we find that

$$\frac{k}{\omega}=1+\delta n+48\pi g\omega^2\delta n^2. \tag{19}$$

In the lowest approximation in the nonlinearity, the expression for δn for the two polarization states mentioned earlier has the form[1]

$$\delta n=K\gamma B_0^2\sin^2\theta,\quad K_\perp=\tfrac{7}{2},\quad K_\|=2, \tag{20}$$

where θ is the angle between $\mathbf{k}$ and $\mathbf{B}_0$.

Equation (19) clearly shows that the wave number is a nonlinear function of frequency, with the nonlinearity coefficient being in turn quadratic in the nonlinear shift δn of the refractive index. Accordingly, for the group velocity v_g of a pulse of electromagnetic radiation we have

$$\frac{1}{v_g}=\frac{dk}{d\omega}=1+\delta n+144\pi g\omega^2\delta n^2. \tag{21}$$

The spread of pulses is characterized by the quadratic dispersion parameter[15]

$$D_2=\frac{d^2k}{d\omega^2}=288\pi g\omega\delta n^2. \tag{22}$$

The inequality $D_2>0$ corresponds to normal (in contrast to anomalous) dispersion, which is natural due to the fact that the system is far from resonance ($\omega\ll\omega_c$). Here the characteristic length of spread of a pulse with an initial length τ_p is

$$L_{\text{disp}}=\frac{\tau_p^2}{D_2}. \tag{23}$$

We see that vacuum dispersion manifests itself most vividly when the high-frequency radiation is in the form of short pulses.

4. SELF-ACTION OF RADIATION

Dispersion is important in various nonlinear optics effects in vacuum, including its effect on the probability of a photon decaying into two photons in an external field,[1] a process that occurs when synchronism conditions are met (in the terminology used in nonlinear optics this is called parametric decay of high-frequency radiation[16]). Below we study the lowest-threshold phenomena of self-action of radiation in a vacuum that are possible in the presence of dispersion. Let us take the simplest variant of the polarization structure of the radiation. We send the z axis along the direction of preferable propagation of the high-frequency radiation and assume that the electric strength vector is directed approximately along the x axis (unit vector $\mathbf{e}_x$) and the magnetic strength vector, along the y axis (unit vector $\mathbf{e}_y$). When both static fields are present, the electric (vector $\mathbf{E}_0=E_0\mathbf{e}_x$) and the magnetic (vector $\mathbf{B}_0=B_0\mathbf{e}_y$), Eqs. (19), (21), and (22) retain their form, but instead of (20) we have[12]

$$\delta n=2\gamma(E_0-B_0)^2. \tag{24}$$

Now, in contrast to Sec. 3, we take into account not only the linear terms but also the lowest-order nonlinear (cubic) terms in the alternating fields. Here we employ the standard approximation of nonlinear optics, the approximation of slowly varying amplitudes, or the quasioptical equation method,[16] i.e., we assume that the envelope E_1 of the alternating field changes little in the course of approximately one period and over distances of the order of one wavelength of the high-frequency vibrations. To this end we write the strength of the alternating field in the form

$$E'=\tfrac{1}{2}E_1\exp[i\omega(z-t)]+\text{c.c.} \tag{25}$$

If we ignore the nonresonant terms (which do not meet the synchronism conditions), we arrive at the following quasioptical equation:

$$\frac{\partial E_1}{\partial z}+\frac{1}{v_g}\frac{\partial E_1}{\partial t}+\frac{1}{2i\omega}\Delta_\perp E_1+i\frac{D_2}{2}\frac{\partial^2E_1}{\partial t^2}-i\omega n_2|E_1|^2E_1=0, \tag{26}$$

where

$$\Delta_\perp=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2},\quad n_2=\frac{e^4}{360\pi m^4}\delta n^2. \tag{27}$$

The nonlinearity coefficient n_2 is positive, which corresponds to self-focusing (in contrast to defocusing). Note that in the given approximation the nonlinearity coefficient is frequency-independent, which means that there is no nonlinearity dispersion.

The quasioptical equation (26) describes the effects of diffraction, dispersion, and nonlinearity of relatively low-frequency electromagnetic radiation in vacuum. There is reason to study the manifestations of these factors separately, using the analogy between Eq. (26) and the quasioptical equation for an ‘‘ordinary’’ Kerr medium.[15–18] What is important here is that the dispersion and nonlinearity coefficients have the same sign ($D_2>0$ and $n_2>0$).

We start with the case of continuous or quasicontinuous radiation (the pulse length τ_p is so large that dispersion distortions along the length L of the path have no effect, $L\ll L_{\text{disp}}$). Then in (26) we can ignore the term containing D_2. If we shift to a reference frame that moves with a velocity equal to that of the group velocity of the radiation, v_g, with t replaced by

$$\tau=t-z/v_g, \tag{28}$$

Eq. (26) transforms into the following nonlinear Schrödinger equation:

$$i\omega\frac{\partial E_1}{\partial z}+\frac{1}{2}\Delta_\perp E_1+\omega^2n_2|E_1|^2E_1=0. \tag{29}$$

This equation is well known from the literature (see Refs. 16 and 17). It describes the balance between the diffraction spread of a beam of light and the beam's nonlinear collapse, i.e., the self-focusing and collapse in a Kerr medium of radiation whose power exceeds the critical value P_{sf}. Here

$$P_{\mathrm{sf}}=\frac{5.763}{8\pi n_2\omega^2}=\frac{5.763c}{8\pi n_2k^2}. \tag{30}$$

In this expression for the critical self-focusing power we used ordinary units and the wave vector $k=\omega/c$. In terms of these units the nonlinearity coefficient n_2 is

$$n_2=\frac{\alpha^3}{475\pi^3}\left(\frac{E_0-B_0}{E_c}\right)^4\frac{1}{E_c^2}, \tag{31}$$

where $\alpha=e^2/\hbar c=1/137$ is the fine structure constant. If the radiation power exceeds the critical value, within the framework of the quasioptical equation (29) the radiation becomes focused, with the intensity near a nonlinear focal point tending to infinity (collapse). Actually, the intensity is not infinite at the focal point because of factors that are not taken into account by the quasioptical equation: absorption of radiation in the vacuum due to intense electron–positron pair production (''vacuum breakdown''), the effect of higher-order nonlinearity terms in the field's Lagrangian, and the decrease of the width of the beam in focusing to values of order of the radiation's wavelength. Note that another variant of self-focusing, i.e., small-scale self-focusing, or filamentation, in which the beam is split into separate intense filaments,[19] requires considerably higher radiation powers ($P\gg P_{\mathrm{sf}}$). In accordance with (30), the critical power decreases with increasing radiation frequency as ω^{-2}. Even if the radiation power is lower than the critical value, vacuum nonlinearity affects primarily the angular divergence of the beam, with the characteristic length of nonlinear distortions being

$$L_{\mathrm{nl}}=\frac{1}{kn_2E_c^2}\frac{E_c^2}{|E|^2}.$$

Another important example of self-action is related to the balance between the dispersion spread and the nonlinear squeezing of radiation pulses.[15] In this case, assuming that the path length is smaller than the characteristic length of transverse distortions of the beam, we can ignore the term with the Laplace operator in (26) and use (28):

$$i\frac{\partial E_1}{\partial z}-\frac{D_2}{2}\frac{\partial^2E_1}{\partial\tau^2}+\omega n_2|E_1|^2E_1=0. \tag{32}$$

This one-dimensional Schrödinger equation allows for an exact solution by the inverse scattering method.[20] Soliton-like solutions of Eq. (32), in which the intensity profile of the pulse does not change in the course of pulse propagation, play a special role. Since both D_2 and n_2 are positive, this equation has no solution of the ''bright-soliton'' type (a pulse with zero intensities at the leading and trailing edges). However, Eq. (32) has solutions of the ''dark-soliton'' type (a pulse with a constant background at the periphery and an intensity gap at the center; see, e.g., Refs. 15 and 16):

$$E_1=A\ \tanh(q\tau)\exp(i\Gamma z). \tag{33}$$

The soliton parameters, i.e., the amplitude A at the periphery, the reciprocal pulse length q, and the propagation constant Γ, are linked by the following relationship:

$$\Gamma=D_2q^2=\omega n_2A^2. \tag{34}$$

A dark soliton can be interpreted as the propagation of a shock wave envelope whose front (the difference between the two values of the envelope amplitude, A and A) is stabilized due to dispersion. For numerical estimates it is convenient to express the frequency and strength of the electromagnetic field in terms of the corresponding critical values given above: $\omega_c/2\pi=6.2\times10^{18}$ Hz and $E_c=4.4\times10^{13}$ esu. Then (22) becomes

$$D_2\left[\frac{\mathrm{s}^2}{\mathrm{cm}}\right]=3.34\times10^{43}K^2\frac{\omega}{\omega_c}\left(\frac{B_0}{E_c}\sin\theta\right)^4, \tag{35}$$

and the nonlinearity coefficient (31) becomes

$$n_2[\mathrm{esu}]=1.4\times10^{-38}\left(\frac{E_0-B_0}{E_c}\right)^4. \tag{36}$$

If in accordance with the conditions (5) we put $\omega/\omega_c=0.1$ and $(E_0-B_0)/E_c=0.1$, we have $D_2=10^{-47}$ s^2/cm and $n_2=10^{-42}$ esu.

Thus, we have formulated quasioptical equations that describe the propagation of narrow and quasimonochromatic beams and pulses of relatively low-frequency and weak electromagnetic radiation (conditions (5)) in an electron–positron vacuum in the presence of static electric and magnetic fields. The allowance for vacuum diffraction, dispersion, and nonlinearity has made it possible, by using the analogy with effects of self-focusing of radiation in a Kerr medium, to demonstrate the possibility of self-focusing and the formation of dark solitons in extreme astrophysical conditions. Such effects can also manifest themselves in experiments with ultrashort laser pulses in the crossed beams geometry.

The author is grateful to S. V. Fedorov for fruitful discussions. The work was sponsored by the Russian State Committee for Higher Education (Grant No. 95-0-5.5-74 for research in the field of basic science).

[1] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 3rd ed., Pergamon Press, Oxford (1991).

[2] M. Lutzky and J. S. Toll, Phys. Rev. **113**, 1649 (1959).

[3] T. Eber, Rev. Mod. Phys. **38**, 626 (1966).

[4] S. L. Adler, Ann. Phys. (N.Y.) **67**, 599 (1971).

[5] A. E. Shabad, Ann. Phys. (N.Y.) **90**, 166 (1975).

[6] Z. Bialynicka-Birula, Physica **20**, 513 (1981).

[7] E. B. Aleksandrov, A. A. Ansel'm, and A. N. Moskalev, Zh. Éksp. Teor. Fiz. **89**, 1181 (1985) [Sov. Phys. JETP **62**, 680 (1985)].

[8] N. N. Rozanov, Zh. Éksp. Teor. Fiz. **103**, 1996 (1993) [JETP **76**, 991 (1993)].

[9] N. N. Rozanov, Proc. SPIE (1993), Vol. 2097, p. 252.

[10] S. G. Mamaev, V. M. Mostepanenko, and M. I. Éĭdes, Yad. Fiz. **33**, 1675 (1981) [Sov. J. Nucl. Phys. **33**, 569 (1981)].

[11] A. A. Grib, S. G. Mamaev, and V. M. Mostepanenko, *Vacuum Quantum Effects in Intense Fields* [in Russian], Atomizdat, Moscow (1988).

[12] N. N. Rozanov and S. V. Fedorov, Opt. Spektrosk. (1998) (accepted for publication).

[13] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed., Pergamon Press, Oxford (1975).

[14] G. G. Pavlov and Yu. A. Shibanov, Zh. Éksp. Teor. Fiz. **76**, 1457 (1979) [Sov. Phys. JETP **49**, 741 (1979)].

[15] S. A. Akhmanov, V. A. Vysloukh, and A. S. Chirkin, *The Optics of Femtosecond Laser Pulses* [in Russian], Nauka, Moscow (1988).

[16] M. V. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *Theory of Waves* [in Russian], Nauka, Moscow (1990).

[17] V. N. Lugovoĭ and A. M. Prokhorov, Usp. Fiz. Nauk **111**, 203 (1973) [Sov. Phys. Usp. **16**, 658 (1974)].

[18] N. N. Rozanov, in *Progress in Optics*, E. Wolf (ed.), Vol. 35, North-Holland, Amsterdam (1996), p. 1.

[19] V. I. Bespalov and V. I. Talanov, Pis'ma Zh. Éksp. Teor. Fiz. **3**, 471 (1966).

[20] S. V. Manakov, S. P. Novikov, L. P. Pitaevskiĭ, and V. E. Zakharov, *Theory of Solitons: The Inverse Scattering Method*, Consultants Bureau, New York (1984).

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Selective reflection of resonance radiation from excited media

B. A. Veklenko, R. B. Gusarov, and Yu. B. Sherkunov

Moscow Power Engineering Institute, 111250 Moscow, Russia

(Submitted 19 April 1997)
Zh. Éksp. Teor. Fiz. **113**, 521–538 (February 1998)

According to quantum electrodynamics, the cross section for resonant scattering of radiation on an aggregate of excited atoms can be written as a sum of positive definite terms. This type of structure is not consistent with the Fresnel formulas for the reflection coefficient of radiation from thermally excited media. The difference shows up on a macroscopic level and indicates that semiclassical radiation theory cannot be used. A study of the correlation between elastic scattering and stimulated emission processes clarifies the reason for the discrepancies. The resulting singularities require summing of Feynman diagrams which appear beginning in the sixth order of perturbation theory. A lower bound estimate for the reflection coefficient from a plane layer is given, including processes which violate the statistics of radiation. The contribution of stimulated emission processes caused by the initially scattered photon are examined specifically. An experiment is proposed which would settle the choice of theories. © *1998 American Institute of Physics.* [S1063-7761(98)01002-6]

1. INTRODUCTION

Research on the kinetics of intense electromagnetic fields in matter has required a generalization of transport theory for resonance radiation[1–3] to the case of excited media. This kind of generalization has been carried out in a number of papers.[4–10] The most general results have been obtained based on a method employing quantum mechanical kinetic Green functions.[9–11] While this generalization is entirely suitable for uniform or quasiuniform media, when inhomogeneities induce reflection effects or when the boundary of the radiation diffusion region is taken into account, new unaccounted-for physical phenomena arise.[12,13] In fact, the derivation of a closed system of equations in the standard quantum mechanical Green function technique requires a discontinuity in the photon–photon correlation functions, which raises doubt as to whether stimulated emission processes have been correctly taken into account. Let $\hat{\alpha}_{\mathbf{k}\lambda}$ be the photon annihilation operator in the mode defined by the wave vector $\mathbf{k}$ and polarization λ. Here we are concerned with a correct accounting for the correlators $\langle \hat{\alpha}^{+}_{\mathbf{k}\lambda} \hat{\alpha}^{+}_{\mathbf{k}\lambda} \hat{\alpha}_{\mathbf{k}\lambda} \hat{\alpha}_{\mathbf{k}\lambda} \rangle$, where all the indices $(\mathbf{k},\lambda)$ are the same. In the standard technique for one-particle Green functions, correlators of this type are neglected, citing their small statistical weight.[14] In thermodynamic equilibrium in the thermodynamic limit, it can be shown that this neglect is justified.[14] In calculating kinetic phenomena, it must be borne in mind that a description of each stimulated emission event resulting from a resonance process will involve the appearance of a Dirac δ-function. Thus, these terms are singular and their neglect becomes dubious. The use of perturbation theory as a whole is dubious for the same reason.

The necessity of a correct accounting for the photon–photon correlators shows up distinctly in the phenomena of selective reflection of resonance radiation from vacuum–excited medium interfaces. This type of boundary can be regarded as a limiting case of spatial inhomogeneity. The selective reflection of resonance radiation from excited media has been studied both experimentally[15–18] and theoretically.[18–24] Despite some optimistic claims,[18] there is no agreement between theory and experiment and even among different theoretical calculations. Most of the calculations have been based on semiclassical radiation theory, operating with an unquantized electromagnetic field. In that case, the calculations yield the Fresnel formulas for the reflection coefficient R when there is no inverted population in the medium. At low concentrations of scattering particles, the coefficient R, which is determined by the refractive index of the medium, is proportional[25] to the square of the difference of the concentrations of the scatterers in the ground and excited states, i.e., $R \propto (n_1 - n_2)^2$. The same result follows from quantum electrodynamics when there is a discontinuity in the photon–photon correlators. It has been shown[12] that a proper quantum electrodynamical treatment of these correlators for resonance radiation changes the result fundamentally. It turns out that the reflection coefficient cannot be expressed in terms of the refractive index of the medium, and the estimated lower bound is proportional to the total density of scattering atoms, i.e. $R \propto (n_1 + n_2)^2$.

In this paper our goal is to find the mechanism for this behavior, which at first glance seems so unusual, by studying the elementary scattering and stimulated emission events, as well as the correlation between them. Our goal is to clarify the structure of the singular terms and point out the reason that perturbation theory is unacceptable. Thus, we are concerned with particular properties of the diffusion of resonance radiation in excited media. The mathematical problem is essentially how to sum the infinite series. It is perfectly clear that under steady state conditions, the scattering of resonance radiation cannot be studied using perturbation theory. Usually, in accordance with the Weisskopf–Wigner theory,[26] it is assumed that the difficulty is removed when the finite width of the energy levels of the atoms is taken into account. In the following it will be explained why this procedure for summing the Feynman diagrams is not adequate for the theory of resonance radiation transport in excited media, and another consistent procedure for summing the diagrams will be pointed out.

2. QUALITATIVE ANALYSIS OF THE PROBLEM

Some properties of the reflection coefficient can be obtained from general considerations. We denote the eigenfunctions of the atoms in a reflecting medium by φ_i. We take note of the adiabatic hypothesis[27] and assume that the atoms were in state φ_0 before scattering. The complete wave function of the medium and the electromagnetic field Ψ can be expanded in φ_i:

$$\Psi = f_0\varphi_0 + \sum_{i\neq 0} f_i\varphi_i. \tag{1}$$

If we are interested in the density matrix ρ of the electromagnetic field, then it is necessary to form a bilinear construct from Eq. (1) which, in turn, must be summed over the arguments of the atoms in the medium. We find

$$\rho = \rho^{(c)} + \rho^{(n)},$$

where

$$\rho^{(c)} = \mathrm{Tr}_a\, f_0 f_0^*, \quad \rho^{(n)} = \mathrm{Tr}_a \sum_{i\neq 0} f_i f_i^*. \tag{2}$$

The matrix $\rho^{(c)}$ describes the scattered field when the atoms of the medium do not change state (coherent scattering channel). The matrix $\rho^{(n)}$ corresponds to scattering processes in which the initial state of the atoms in the medium changes (incoherent channel). The incoherent channel includes spontaneous emission and Raman scattering, as well as stimulated atomic emission. Note that stimulated emission and coherent scattering processes are described by different channels in a rigorous quantum mechanical theory. Since the different channels correspond to different final-state wave functions, they do not interfere with one another.

Suppose we are interested in the average occupation numbers of phonons in mode ($\mathbf{k}$,λ) owing to scattering, i.e.,

$$\langle N_{\mathbf{k}\lambda}\rangle = \mathrm{Tr}_{ph}\, \hat{N}_{\mathbf{k}\lambda}\rho.$$

Here a sum is taken over the arguments of the photon field and

$$\hat{N}_{\mathbf{k}\lambda} = \hat{\alpha}^+_{\mathbf{k}\lambda}\hat{\alpha}_{\mathbf{k}\lambda}.$$

According to Eq. (2), we have

$$\langle N_{\mathbf{k}\lambda}\rangle = \langle N_{\mathbf{k}\lambda}\rangle^{(c)} + \langle N_{\mathbf{k}\lambda}\rangle^{(n)},$$

with

$$\langle N_{\mathbf{k}\lambda}\rangle^{(c)} = \mathrm{Tr}_{ph}\ \mathrm{Tr}_a |\hat{\alpha}_{\mathbf{k}\lambda} f_0|^2 > 0,$$

$$\langle N_{\mathbf{k}\lambda}\rangle^{(n)} = \mathrm{Tr}_{ph}\ \mathrm{Tr}_a \sum_{i\neq 0} |\hat{\alpha}_{\mathbf{k}\lambda} f_i|^2 > 0.$$

In a study[24] of the coherent channel for reflection of photons from an interface boundary, it is shown that $\langle N_{\mathbf{k}\lambda}\rangle^{(c)} \sim (n_1+n_2)^2$, implying that this is an estimated lower bound for R, since $\langle N_{\mathbf{k}\lambda}\rangle^n > 0$. This statement is consistent with the above analysis, but not the semiclassical theory of radiation. Furthermore, let us expand the coefficients f_i in eigenfunctions of the photon number operator $\hat{N}_{\mathbf{k}\lambda}$:

$$f_i = \sum_N f_{iN}|N\rangle.$$

Now it is obvious that

$$\langle N_{\mathbf{k}\lambda}\rangle^{(n)} = \mathrm{Tr}_{ph}\ \mathrm{Tr}_a \sum_{iN} N|f_{iN}|^2 > 0. \tag{3}$$

Thus, the contribution of the incoherent channel to R can be represented as an expansion in the number of scattered photons, which is analogous to the expansion in the Tamm–Dankov method.[28] Note that all terms in the expansion (3) must be positive.

3. SCATTERING OF LIGHT BY TWO ATOMS

We now analyze the situation from the standpoint of perturbation theory. We consider resonant scattering by two atoms. Let an unperturbed atom lie at the point $\mathbf{R}_1$ and an excited atom, at $\mathbf{R}_2$. Of course, under steady state conditions, perturbation theory based on the $\hat{S}$ matrix is not applicable here. We proceed as follows: we track the $\pm i0$ terms in the resonance denominators that appear owing to the adiabatic hypothesis. Replacing these terms by finite expressions when the energy widths γ of the atoms are taken into account removes the resonance divergences. This kind of replacement takes place without a change in the signs of the $i0$, which would violate causality. Thus, the $i0$ terms make it possible to foresee the general structure of the result in the more complete theory.

Let two atoms with a single valence electron apiece interact with a coherent quantized electromagnetic field. The interaction is assumed to be quasiresonant, with $|k-\omega_{21}| \ll k+\omega_{21}$, which makes it possible to neglect the square of the vector potential in the total Hamiltonian of the system. Here ω_{21} denotes the frequency of the resonant transition in the atoms, and $\hbar = c = 1$. In the second quantization representation, we take the Hamiltonian of the system to be

$$\hat{H} = \hat{H}^0 + \hat{H}',$$

$$\hat{H}^0 = \sum_j \varepsilon'_j \hat{b}^+_j \hat{b}_j + \sum_j \varepsilon_j \hat{\beta}^+_j \hat{\beta}_j + \sum_{\mathbf{k}\lambda} k\left(\hat{\alpha}^+_{\mathbf{k}\lambda}\hat{\alpha}_{\mathbf{k}\lambda} + \frac{1}{2}\right),$$

$$\hat{H}' = -\frac{e}{m}\int \hat{\psi}^+(\mathbf{r})\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}(\mathbf{r})\hat{\psi}(\mathbf{r})\, d\mathbf{r} - \frac{e}{m} \times \int \hat{\varphi}^+(\mathbf{r})\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}(\mathbf{r})\hat{\varphi}(\mathbf{r}) d\mathbf{r}.$$

Here ε'_j and ε_j are the allowed energies of the first and second atoms, and $\hat{b}_j(\hat{b}^+_j)$ and $\hat{\beta}_j(\hat{\beta}^+_j)$ are the annihilation (creation) operators of the valence electron states corresponding to these energy levels. We neglect spin and exchange effects. Under these conditions, for simplicity, all operators can be regarded as Bose-type.

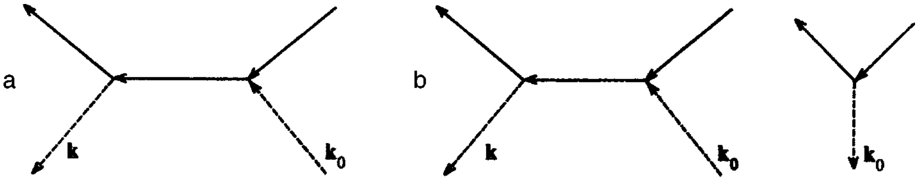


FIG. 1. Feynman diagrams for the coherent (a) and incoherent (b) scattering channels.

The field operators for the first and second atoms correspond to the expressions

$$\hat\psi(\mathbf{r})=\sum_j \psi_j(\mathbf{r}-\mathbf{R}_1)\hat b_j,\quad \hat\varphi(\mathbf{r})=\sum_j \varphi_j(\mathbf{r}-\mathbf{R}_2)\hat\beta_j,$$

where ψ_j and φ_j are the wave functions of the electronic states in the atoms. The electromagnetic field operator is

$$\hat{\mathbf{A}}(\mathbf{r})=\sum_{\mathbf{k}\lambda}\frac{\mathbf{e}_\mathbf{k}^\lambda}{\sqrt{2kV}}(\hat\alpha_{\mathbf{k}\lambda}e^{i\mathbf{k}\cdot\mathbf{r}}+\hat\alpha_{\mathbf{k}\lambda}^+e^{-i\mathbf{k}\cdot\mathbf{r}}),$$

where $\mathbf{e}_\mathbf{k}^\lambda$ are the unit vectors of the linear polarization and $V=L_xL_yL_z$ is the quantization volume.

In the Schrödinger picture, the wave equation of the system has the form

$$i\frac{\partial\Psi}{\partial t}=(\hat H^0+\hat H')\Psi.$$

The transition to the interaction picture is carried out in the standard way using the operator $\exp(i\hat H^0 t)$, whereupon

$$i\frac{\partial\Psi}{\partial t}=\hat H'\Psi$$

and

$$\Psi(t)=\hat S(t,-\infty)\Psi^0,$$

$$\hat S(t,-\infty)=\hat T\exp\left\{-i\int_{-\infty}^t\hat H'(t)dt\right\}.$$

For a preliminary analysis we limit ourselves to third-order perturbation theory:

$$\hat S=1+\hat S^{(1)}+\hat S^{(2)}+\hat S^{(3)}.$$

The operators $\hat S^{(1)}$ and $\hat S^{(3)}$ can show up only when describing the incoherent scattering channel. If we exclude Raman scattering and two-photon absorption (emission), then the operator $\hat S^{(2)}$ will only describe coherent scattering. Thus,

$$\langle N_{\mathbf{k}\lambda}\rangle=\langle N_{\mathbf{k}\lambda}\rangle^{(c)}+\langle N_{\mathbf{k}\lambda}\rangle^{(n)},$$
$$\langle N_{\mathbf{k}\lambda}\rangle^{(c)}=\langle\hat S^{(2)}\hat N_{\mathbf{k}\lambda}\hat S^{(2)}\rangle, \tag{4}$$
$$\langle N_{\mathbf{k}\lambda}\rangle^{(n)}=\langle\hat S^{(1)}+\hat S^{(3)}|\hat N_{\mathbf{k}\lambda}|\hat S^{(1)}+\hat S^{(3)}\rangle.$$

Averaging is over the initial state of the system. As might be expected, both terms on the right hand side of the first of Eqs. (4) are positive definite.

We begin with the coherent channel, omitting the details in the standard computational procedure. We merely note that the coupling of atomic operators that arises according to the Feynman diagram (Fig. 1a) corresponds to the following analytic expression:

$$\overline{\hat\psi(\mathbf{r},t)\hat\psi^+(\mathbf{r}',t')}$$
$$=i\sum_j\int\exp[-iE(t-t')]\frac{\psi_j(\mathbf{r})\psi_j^*(\mathbf{r}')}{E-\varepsilon_j+i0}\frac{dE}{2\pi}.$$

An analogous expression corresponds to the coupling of the operators $\hat\varphi$ and $\hat\varphi^+$. It is the ''+'' sign in front of the $i0$ that shows up in the final result, which has the form

$$\langle N_{\mathbf{k}\lambda}\rangle^{(c)}=4\pi^2\left|\frac{P_{21}^\lambda(\mathbf{k})}{\sqrt{2kV}}\right|^2\left|\frac{P_{21}^{\lambda_0}(\mathbf{k}_0)}{\sqrt{2k_0V}}\right|^2\delta^2(k-k_0)$$
$$\times\left|\frac{\exp[i(\mathbf{k}_0-\mathbf{k})\cdot\mathbf{R}_1]}{k-\omega_{21}+i0}+\frac{\exp[i(\mathbf{k}_0-\mathbf{k})\cdot\mathbf{R}_2]}{\omega_{21}-k+i0}\right|^2N_0. \tag{5}$$

We have limited ourselves to a two-level approximation for the atoms, which are assumed to have identical structure,

$$P_{ij}^\lambda(\mathbf{k})=\frac{e}{m}\int\psi_i^*(\mathbf{r})\hat{\mathbf{p}}\cdot\mathbf{e}_\mathbf{k}^\lambda e^{i\mathbf{k}\cdot\mathbf{r}}\psi_j(\mathbf{r})\,d\mathbf{r};$$

N_0 is the number of photons in the scattering mode.

The presence of δ^2 in the result should not be confusing, since this kind of dependence is typical of quantum electrodynamics, and makes it possible to calculate the probabilities of processes per unit time. Note a characteristic feature of Eq. (5). A term of the form $\sin^2(\mathbf{k}_0-\mathbf{k})\cdot(\mathbf{R}_1-\mathbf{R}_2)$ appears here for nonresonant frequencies $(k_0\neq\omega_{21})$. For the resonance frequencies $(k_0=\omega_{21})$, the dependence is different and a factor $\cos^2(\mathbf{k}_0-\mathbf{k})\cdot(\mathbf{R}_1-\mathbf{R}_2)$ shows up.

In the following we shall be interested in the reflection coefficient at the interface. A specularly reflected beam corresponds to a geometry in which $(\mathbf{k}_0-\mathbf{k})\perp(\mathbf{R}_1-\mathbf{R}_2)$. It is now obvious that for the nonresonance frequencies the result is zero according to Eq. (5), and that for Fresnel reflection from an excited medium, this corresponds to the dependence

$$\langle N_{\mathbf{k}\lambda}\rangle^{(c)}\sim(n_1-n_2)^2N_0,\quad k_0\neq\omega_{21}. \tag{6}$$

At the resonance frequencies, for the same reason, instead of Eq. (6) it turns out that

$$\langle N_{\mathbf{k}\lambda}\rangle^{(c)}\sim(n_1+n_2)^2N_0,\quad k_0=\omega_{21}. \tag{7}$$

It is too early to carry out a complete analysis including the incoherent scattering channel. But, since the coherent scattering channel defines a lower bound on the reflection coefficient, and Eq. (7) divulges the existence of a reflected wave even for an optically saturated medium ($n_1=n_2$), it is evident that the Fresnel formula breaks down. Thus, the reflection of resonance radiation from excited media cannot be calculated using semiclassical radiation theory, which yields the Fresnel formula.

Let us consider the incoherent channel. The term $\langle\hat{S}^{(1)}\hat{N}_{\mathbf{k}\lambda}\hat{S}^{(1)}\rangle$ is of no interest to us, since it describes absorption or stimulated emission by the atoms. Now, according to Eq. (4), we have to calculate $\hat{S}^{(3)}$ in the lowest order of perturbation theory. The scattering process, which contains the interference component responsible for selective reflection, corresponds to the Feynman diagram in Fig. 1b. Its contribution to $\langle N_{\mathbf{k}\lambda}\rangle^{(n)}$ is

$$\begin{aligned}\langle N_{\mathbf{k}\lambda}\rangle^{(n)}&=\langle\hat{S}^{(1)}\hat{N}_{\mathbf{k}\lambda}\hat{S}^{(3)}\rangle+\text{c.c.}\\&=8\pi^3 i\left|\frac{P_{21}^{\lambda}(\mathbf{k})}{\sqrt{2kV}}\right|^2\left|\frac{P_{21}^{\lambda_0}(\mathbf{k}_0)}{\sqrt{2k_0V}}\right|^2\\&\quad\times\frac{\delta^2(k-k_0)\,\delta(\omega_{21}-k)}{k-\omega_{21}-i0}\\&\quad\times\exp[-i(\mathbf{k}_0-\mathbf{k})\cdot(\mathbf{R}_2-\mathbf{R}_1)]+\text{c.c.}\end{aligned}\tag{8}$$

This result is remarkable in many regards. First and foremost, its sum with the coherent channel yields

$$\begin{aligned}\langle N_{\mathbf{k}\lambda}\rangle^{(c)}+\langle N_{\mathbf{k}\lambda}\rangle^{(n)}&=4\pi^2\left|\frac{P_{21}^{\lambda}(\mathbf{k})}{\sqrt{2kV}}\right|^2\left|\frac{P_{21}^{\lambda_0}(\mathbf{k}_0)}{\sqrt{2k_0V}}\right|^2\\&\quad\times\left|\frac{\exp[i(\mathbf{k}_0-\mathbf{k})\cdot\mathbf{R}_1]}{k-\omega_{21}+i0}\right.\\&\quad\left.+\frac{\exp[i(\mathbf{k}_0-\mathbf{k})\cdot\mathbf{R}_2]}{\omega_{21}-k-i0}\right|^2N_0\delta^2(k-k_0),\end{aligned}\tag{9}$$

which differs from Eq. (5) only in the sign $i0$ in the second term. This difference is crucial. Above all, Eq. (9) is a complete result from fourth-order perturbation theory. When the Heisenberg picture is used in the calculations, this result shows up immediately, since there is no separation into coherent and incoherent channels there.

This same result also follows from the semiclassical theory of scattering. According to Eq. (9), there is never any scattering in the direction of the vector $\mathbf{k}$ corresponding to the condition $(\mathbf{k}_0-\mathbf{k})\perp(\mathbf{R}_1-\mathbf{R}_2)$. This implies that Eq. (6), along with the Fresnel equations, are valid at arbitrary frequencies k_0. This conclusion contradicts the above analysis. The error lies in Eq. (8), which yields an estimate at the resonance frequencies of

$$\langle N_{\mathbf{k}\lambda}\rangle^{(n)}\sim-\cos[(\mathbf{k}_0-\mathbf{k})\cdot(\mathbf{R}_1-\mathbf{R}_2)]<0,$$

which is inconsistent with the positive definiteness of the coherent scattering channel. Therefore, Eq. (9), which follows from the lowest-order scattering theory, is false, along with Eq. (8). The situation is not saved by taking the finite width of the energy levels into account, which does not change the sign of $\langle N_{\mathbf{k}\lambda}\rangle^{(n)}$.

The positive definiteness of the incoherent channel can be recovered if similar terms in the expansion of the $\hat{S}$ matrix are included in the two functions of the scalar product in Eq. (4). This means that systematic allowance for $\hat{S}^{(3)}$ requires that perturbation theory terms of sixth order be retained in Eq. (4). This implies, in turn, that the sixth-order terms are larger in absolute value than the fourth-order result and contain singularities, which means that perturbation theory is not applicable on the whole. Including the finite values of γ does not remedy the situation. Thus, with finite γ, the sixth-order terms still contain singularities whose elimination requires summation of infinite subsequences of Feynman diagrams.

It is necessary to go beyond the confines of fourth-order perturbation theory. Even sixth-order perturbation theory introduces qualitatively new phenomena. It is natural to begin a study of the contribution from sixth-order perturbation theory with an analysis of the processes that bring a single photon to the scattering mode. Here we encounter the square of the same element $\hat{S}^{(3)}$ corresponding to the Feynman diagrams of Fig. 1b. The result is nonlinear in N_0. We shall not examine these processes. Processes that bring two photons to the scattering mode in sixth-order perturbation theory include stimulated emission from an excited atom, induced by photons initially scattered by another atom. This coupling of processes is linear in N_0. Including this coupling, which corresponds to terms of the type $|f_{i2}|^2$ in Eq. (3), along with the coherent channel, only improves the lower bound estimate of the overall result for R. The physics of these processes is important for predicting the results of reflection from media with a population inversion.

In the following, we examine the reflection of light from media with a plane-parallel configuration. Coherent scattering will be taken into account, along with the two-photon reflection processes discussed above. Two-photon reflection is calculated by summing infinite subsequences of Feynman diagrams.

4. REFLECTION OF LIGHT FROM A SEMI-INFINITE MEDIUM

We assume that the half-space $z>0$ is filled with an atomic gas consisting of atoms with a single valence electron. Radiation in the $(\mathbf{k}_0,\lambda_0)$ mode is incident on the interface boundary. We take a Schrödinger equation of the form

$$i\,\frac{\partial\Psi}{\partial t}=\left[\hat{H}_a+\hat{H}_{ph}-\frac{e}{m}\int\hat{\psi}^{+}(\mathbf{r},\mathbf{R})\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}(\mathbf{r})\hat{\psi}(\mathbf{r},\mathbf{R})\,d\mathbf{r}\,d\mathbf{R}\right]\Psi,$$

where

$$\hat{H}_a=\sum_{i\mathbf{p}}\varepsilon_i(p)\hat{b}_{i\mathbf{p}}^{+}\hat{b}_{i\mathbf{p}},\quad \varepsilon_i(p)=\frac{p^2}{2M}+\varepsilon_i,$$

$$\hat{H}_{ph}=\sum_{\mathbf{k}\lambda}k\left(\hat{\alpha}_{\mathbf{k}\lambda}^{+}\hat{\alpha}_{\mathbf{k}\lambda}+\frac{1}{2}\right).$$

ε_i is the energy of an atom in the ith energy state, and $\mathbf{p}$ and M are its momentum and mass. $\hat{b}_{i\mathbf{p}}(\hat{b}_{i\mathbf{p}}^{+})$ denote the Bose operators for annihilation (creation) of atoms in the state $(i,\mathbf{p})$. The gas is assumed to temperature nondegenerate. In addition,

$$\hat{\psi}(\mathbf{r},\mathbf{R})=\sum_{i\mathbf{p}}\psi_i(\mathbf{r}-\mathbf{R})\frac{e^{i\mathbf{p}\cdot\mathbf{R}}}{\sqrt{V}}\hat{b}_{i\mathbf{p}},$$

where $\psi_i(\mathbf{r}-\mathbf{R})$ is the wave function of the ith state of the atom.

To calculate the reflection coefficient, we use the Γ-operator method,[24] which makes it possible to separate the coherent and incoherent scattering channels. We denote the set of occupation numbers for the free electromagnetic field by the vector $\mathbf{N}=\ldots, N_{\mathbf{k}\lambda},\ldots$. In occupation number space this state corresponds to a wave function that can be written in the form

$$\Phi^0(\mathbf{N}|\boldsymbol{\zeta})=\prod_{\mathbf{k}\lambda}\varphi(N_{\mathbf{k}\lambda}|\zeta_{\mathbf{k}\lambda}), \tag{10}$$

where $\varphi(N|\zeta)$ is the quantum mechanical oscillator wave function. The desired Γ-space is constructed as follows. We introduce a creation vector $\rangle_\Gamma^0$, the mathematical vacuum, into this space. Let $\hat{\mathfrak{U}}(\mathbf{N})(\hat{\mathfrak{U}}^+(\mathbf{N}))$ be the annihilation (creation) operator for an aggregate of noninteracting photons with the set of occupation numbers $\mathbf{N}$ and $\hat{\mathfrak{U}}^+(\mathbf{N})\rangle_\Gamma^0$ be the wave function of this state. These wave functions form a complete basis for expanding any wave function describing the physical states. There is a unitary transformation between the basis wave vectors introduced in this way and the wave functions (10).[24] Let

$$\hat{\Phi}(\boldsymbol{\zeta})=\sum_{\mathbf{N}}\hat{\mathfrak{U}}(\mathbf{N})\Phi^0(\mathbf{N}|\boldsymbol{\zeta}).$$

The unitary coupling between the standard occupation number space and the Γ-space is realized through the unitary operator[24]

$$\hat{O}=\hat{\Phi}^+\rangle_\Gamma^0,\quad \Psi_\Gamma=\hat{O}\Psi.$$

In the Γ-space the Schrödinger equation takes the form

$$i\frac{\partial\Psi_\Gamma}{\partial t}=\left[\hat{H}_a+\sum_{\mathbf{N}}\varepsilon(\mathbf{N})\hat{\mathfrak{U}}^+(\mathbf{N})\hat{\mathfrak{U}}(\mathbf{N})-\frac{e}{m}\right.$$
$$\left.\times\int\hat{\Phi}^+\hat{\psi}^+\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}(\mathbf{r})\hat{\psi}\hat{\Phi}\,d\mathbf{r}\,d\mathbf{R}\,d\boldsymbol{\zeta}\right]\Psi_\Gamma,$$

where

$$\varepsilon(\mathbf{N})=\sum_{\mathbf{k}\lambda}k\left(N_{\mathbf{k}\lambda}+\frac{1}{2}\right),\quad d\boldsymbol{\zeta}=\prod_{\mathbf{k}\lambda}d\zeta_{\mathbf{k}\lambda}.$$

The average value of any electromagnetic field operator $\hat{K}$ can be found using the formula

$$\langle K\rangle=\int\langle\hat{\Phi}^+(\boldsymbol{\zeta})\hat{K}\hat{\Phi}(\boldsymbol{\zeta})\rangle_\Gamma\,d\boldsymbol{\zeta},$$

where $\rangle_\Gamma=\Psi_\Gamma$. Thus, the construction

$$\rho(\boldsymbol{\zeta},\boldsymbol{\zeta}')=\langle\hat{\Phi}^+(\boldsymbol{\zeta}')\hat{\Phi}(\boldsymbol{\zeta})\rangle_\Gamma$$

serves as the density matrix of the electromagnetic field in the medium. It can be conveniently calculated using the quantum mechanical Green function formalism[24] in a form proposed by Keldysh[11] in the Γ-space:

$$\mathscr{D}_{ll'}(\boldsymbol{\zeta},t,\boldsymbol{\zeta}',t')=-i\langle\hat{T}_c\check{\Phi}_l(\boldsymbol{\zeta},t)\check{\Phi}^+_{l'}(\boldsymbol{\zeta}',t')\rangle_\Gamma$$
$$=-i\langle\hat{T}_c\tilde{\Phi}_l(\boldsymbol{\zeta},t)\tilde{\Phi}^+_{l'}(\boldsymbol{\zeta}',t')\hat{S}_c\rangle_{0\Gamma}. \tag{11}$$

Here the $\check{\Phi}$ are the field operators in the Heisenberg picture and the $\tilde{\Phi}$ are the field operators in the interaction picture in the Γ-space. The transformation in Eq. (11) from one picture to the other is standard.[24] The subscript l corresponds to a time profile originating $(l=1)$ at the point $t\to-\infty$, extending to $t\to\infty$, and returning again $(l=2)$ to $t\to-\infty$. $\hat{T}_c$ is the chronological operator on this profile. The operator $\hat{S}_c$ has the form

$$\hat{S}_c=\hat{T}_c\exp\left\{\sum_l(-1)^{l+1}\frac{ie}{m}\int\tilde{\Phi}_l^+\tilde{\psi}_l^+\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}(\mathbf{r})\right.$$
$$\left.\times\tilde{\psi}_l\tilde{\Phi}_l d\boldsymbol{\zeta}\,d\mathbf{r}\,d\mathbf{R}\,dt\right\},$$

$$\tilde{\psi}=\sum_{i\mathbf{p}}\psi_i(\mathbf{r}-\mathbf{R})\frac{\hat{b}_{i\mathbf{p}}}{\sqrt{V}}\exp[i\mathbf{p}\cdot\mathbf{R}-i\varepsilon_i(p)t],$$

$$\tilde{\Phi}=\sum_{\mathbf{N}}\tilde{\mathfrak{U}}(\mathbf{N})\Phi^0(\mathbf{N}|\boldsymbol{\zeta})\exp[-i\varepsilon(\mathbf{N})t].$$

In the last term of Eq. (11), the average is taken over the initial system of the atom+field system until the interaction between them is turned on adiabatically. In the following, for Eq. (11) we take this to mean statistical averaging over an ensemble of systems, as well as quantum mechanical averaging. In Eq. (11) it is possible to account explicitly for the interaction of the gas atoms with a reservoir. The effect of the reservoir shows up through the mass operators as a broadening of the energy levels of the atoms. Then the sign of the retarded mass operator is determined by causality.

We assume that the ensemble of atoms has a Gaussian distribution until it interacts with the radiation. This makes it possible to use a thermodynamic variant of Wick's theorem[14] in simplifying the average product of the operators ψ. The product of operators $\tilde{\Phi}$ can be simplified exactly.[24] This makes it possible to avoid a divergence of the correlators that comprise these operators, and thereby to account correctly for the stimulated emission processes of interest to us. Diagram techniques yield the following results: above all, the function $\mathscr{D}_{ll'}$ has the structure

$$\mathscr{D}_{ll'}=\Delta_{ll'}-i\rho_{ll'}.$$

To construct the desired density matrix ρ, it is sufficient to know just ρ_{12}, since $\rho=i\mathscr{D}_{12}$ for $t=t'$. The following system of equations emerges for the matrices $\Delta_{ll'}$ and $\rho_{ll'}$ [24]:

$$\rho_{12}=\rho_{12}^{(c)}+\rho_{12}^{(n)}, \tag{12}$$

$$\rho_{12}^{(c)}=(1+\Delta_r\hat{\mathscr{P}}_r)\rho_{12}^0(1+\hat{\mathscr{P}}_a\Delta_a), \tag{13}$$

and

$$\rho_{12}^{(n)}=-\Delta_r\hat{\mathscr{P}}_{12}^{(n)}\Delta_a, \quad \Delta_r=\Delta_r^0+\Delta_r^0\hat{\mathscr{P}}_r\Delta_r. \tag{14}$$

In addition, $\hat{\mathscr{P}}_{ll'}$ and $\hat{\mathscr{P}}_{ll'}^{(n)}$ are the polarization operators that arise in the Green function technique in Γ-space, and

$$\Delta_r=\Delta_{11}, \quad \Delta_a=-\Delta_{22}=\Delta_r^+,$$

$$\hat{\mathscr{P}}_r=\hat{\mathscr{P}}_{11}, \quad \hat{\mathscr{P}}_a=\hat{\mathscr{P}}_{22}=\hat{\mathscr{P}}_r^+.$$

This kind of structure in the system of equations is convenient for further study in the spirit of the comments in Sec. 2. According to Eq. (12), the density matrix of the photon subsystem breaks up into the sum of two components. The coherent component $\rho_{12}^{(c)}$ describes elastic scattering processes in which the atoms of the medium return to their initial (including translational) quantum state. The incoherent component $\rho_{12}^{(n)}$ describes processes in which the atoms of the medium change their quantum state. The following rules for constructing Feynman diagrams can be used to calculate the polarization operators:

1. A solid line with an arrow is identified with a factor $iG_{ll'}^0$. For homogeneous media

$$G_{ll'}^0=\sum_{jj'\mathbf{p}} \psi_j(\mathbf{r}-\mathbf{R})\psi_{j'}^*(\mathbf{r}'-\mathbf{R}')\int G_{ll'}^{0jj'}(\mathbf{p},E)$$
$$\times\exp[i\mathbf{p}\cdot(\mathbf{R}-\mathbf{R}')-iE(t-t')]\,\frac{dE}{2\pi V}.$$

If the scattering atoms of the medium are subject to the action of a reservoir that broadens the atomic energy levels, then[24]

$$\begin{aligned} G_r^{0jj'}(\mathbf{p},E)&=\delta_{jj'}(E-\varepsilon_j(p)+i\gamma_j/2)^{-1},\\ G_{12}^{jj'}(\mathbf{p},E)&=-2\pi i\,\delta_\gamma(E-\varepsilon_j(p))N_j(\mathbf{p})\delta_{jj'}, \end{aligned} \tag{15}$$

where

$$\delta_\gamma(E)=-\frac{1}{2\pi i}\left[\left(E+i\,\frac{\gamma}{2}\right)^{-1}-\left(E-i\,\frac{\gamma}{2}\right)^{-1}\right],$$

and $N_j(\mathbf{p})$ is the occupation number of the atomic states, $G_r=G_{11}-G_{12}$.

2. A dashed line with an arrow is identified with $i\Delta_{ll'}^0$.

3. A wavy line with an arrow is identified with a factor $\rho_{ll'}^0$.

4. A node, denoted by a circle in a diagram, is identified with a factor $(-1)^{l+1}(ie/m)\hat{\mathbf{p}}\cdot\hat{\mathbf{A}}$. Sometimes we denote a node in a diagram by a dot. In that case the factor $(-1)^{l+1}$ is omitted.

5. At each interior node, integration over all **r**, **R**, t, and ζ is understood.

6. At each exterior mode, integration over **r** and **R** is understood.

Since the product $\tilde{\Phi}\tilde{\Phi}$ vanishes, it follows that each Feynman diagram can either not contain $\rho_{ll'}^0$ or contain the function $\rho_{ll'}^0$ to no more than the first power. For example, the Feynman diagram in Fig. 2 corresponds to the simplest polarization operator

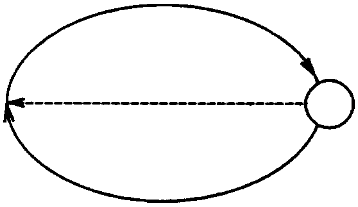


FIG. 2. Feynman diagram for $-i\hat{\mathscr{P}}_{ll'}$ in the Γ-technique.

$$\hat{\mathscr{P}}_{l_1l_2}=-\left(\frac{e}{m}\right)^2\sum_{\nu_1\nu_2}\int \hat{p}^{\nu_1}\hat{A}^{\nu_1}G_{l_1l_2}^0(X_1,X_2)\Delta_{l_1l_2}^0$$
$$\times(-1)^{l_2+1}\hat{p}^{\nu_2}\hat{A}^{\nu_2}G_{l_2l_1}^0(X_2,X_1)\,d\mathbf{r}_1\,d\mathbf{r}_2\,d\mathbf{R}_1\,d\mathbf{R}_2, \tag{16}$$

where $X=\{\mathbf{r},\mathbf{R},t\}$. We restrict ourselves to a two-level approximation for the atoms including degeneracy. We contrast the subscript m for the Zeeman sublevels of their excited state with the subscript μ for the sublevels of the ground state. For uniform filling of the sublevels, after integrating in Eq. (16) with respect to $\mathbf{R}_1$ and $\mathbf{R}_2$, in accordance with Eq. (15) for a uniform infinite medium in the **N**-representation we find

$$\hat{\mathscr{P}}_r(E)=\sum_{\mathbf{k}\lambda}\,[\hat{\alpha}_{\mathbf{k}\lambda}a_r^{\mathbf{k}\lambda}(E-\hat{H}_{ph})\hat{\alpha}_{\mathbf{k}\lambda}^+ +\hat{\alpha}_{\mathbf{k}\lambda}^+c_r^{\mathbf{k}\lambda}$$
$$\times(E-\hat{H}_{ph})\hat{\alpha}_{\mathbf{k}\lambda}],$$

where

$$\begin{aligned} a_r^{\mathbf{k}\lambda}(E)&=\frac{\pi(2j_m+1)\gamma_r}{2\omega_{m\mu}kV}\,\frac{N_m}{E+\omega_{m\mu}+i\gamma/2},\\ c_r^{\mathbf{k}\lambda}(E)&=\frac{\pi(2j_m+1)\gamma_r}{2\omega_{m\mu}kV}\,\frac{N_\mu}{E-\omega_{m\mu}+i\gamma/2}. \end{aligned} \tag{17}$$

Here we have used the sum rule

$$\sum_{m\mu}\,|P_{m\mu}^\lambda(\mathbf{k})|^2=\frac{\pi(2j_m+1)\gamma_r}{\omega_{m\mu}}, \tag{18}$$

where γ_r is the spontaneous emission probability per unit time and j_m is the orbital quantum number. The Doppler effect is left out. When boundaries are present, the polarization operator (16) determines the Poynting vector of the reflected radiation in the coherent channel to lowest approximation in the concentration of scatterers. For a plane-parallel layer of thickness L,[24]

$$\langle\mathbf{s}_\lambda\rangle^{(c)}=\mathbf{k}\,\frac{\sin^2(k_{02}L)}{4k_{0z}^4V}\,|c^{\mathbf{k}\lambda\mathbf{k}_0\lambda_0}(k_0,k_0)|^2N_0,$$

$$\mathbf{k}=\{k_{0x},k_{0y},-k_{0z}\}.$$

Here it is assumed that

$$\rho_{12}^0(E)=2\pi\delta(E-k_0N_0)|N_{\mathbf{k}_0\lambda_0}\rangle\langle N_{\mathbf{k}_0\lambda_0}|.$$

In the linear approximation the reflection coefficient is given by

$$R^{(c)}=\frac{\langle s_\lambda\rangle^{(c)}}{s^0}, \quad s^0=\frac{k_0N_0}{V}, \tag{19}$$

and

$$c^{\mathbf{k}\lambda\mathbf{k}_0\lambda_0}(k_1,k_2)$$

$$=\sum_{m\mu}[P^{\lambda}_{m\mu}(\mathbf{k})]^*P^{\lambda_0}_{m\mu}(\mathbf{k}_0)\left(\frac{n_\mu}{k_1-\omega_{m\mu}+i\gamma/2}\right.$$

$$\left.+\frac{n_m}{\omega_{m\mu}-k_2+i\gamma/2}\right),$$

where

$$n_\mu=\frac{N_\mu}{V}, \quad n_m=\frac{N_m}{V}, \quad \gamma=\gamma_m+\gamma_\mu .$$

The structure of this expression is analogous to that of Eq. (5). In particular, for the resonance frequencies $k=\omega_{m\mu}$,

$$R^{(c)}\sim\left|(n_\mu+n_m)\bar{\lambda}^3\,\frac{\gamma_2}{\gamma}\right|^2<1, \quad \bar{\lambda}=\frac{2\pi}{k}. \tag{20}$$

The last equality serves as a condition for applicability of the theory. For multiple scattering processes, this parameter is raised to the appropriate power. Thus, the coherent scattering channel can be studied by the methods of perturbation theory.

We now study the contribution of the incoherent scattering channel to the Poynting vector. A number of things can be said about this channel. We rewrite the equation for the propagator Δ_r in the form

$$\Delta_r^{-1}=(\Delta_r^0)^{-1}-\hat{\mathscr{P}}_r, \quad (\Delta_r^0)^{-1}=i\,\frac{\partial}{\partial t}-\hat{H}_{ph}. \tag{21}$$

We now apply the operators Δ_r^{-1} and Δ_a^{-1} to the left and right, respectively, of Eqs. (13) and (14). Since the matrix ρ_{12}^0 describes the free electromagnetic field, with the aid of Eq. (21) we find

$$\Delta_r^{-1}\rho_{12}^{(c)}=(\Delta_r^0)^{-1}\rho_{12}^0(1+\hat{\mathscr{P}}_a\Delta_a)=0,$$

$$\Delta_r^{-1}\rho_{12}\Delta_a^{-1}=-\hat{\mathscr{P}}_{12}^{(n)}.$$

Now we operate on the right-hand side of this last equation with the operator Δ_r and on the left with the operator Δ_a. According to Eq. (14), we have

$$\Delta_r\Delta_r^{-1}\rho_{12}\Delta_a^{-1}\Delta_a=\rho_{12}^{(n)}, \tag{22}$$

but

$$\rho_{12}(t,t')=i\mathscr{D}_{12}(t,t')=\langle\check{\Phi}^+\check{\Phi}\rangle_\Gamma. \tag{23}$$

Substituting Eq. (23) into Eq. (22) shows that

$$\rho_{12}^{(n)}=\langle\check{Y}^+(\boldsymbol{\zeta}',t')\check{Y}(\boldsymbol{\zeta},t)\rangle_\Gamma,$$

where

$$\check{Y}=\Delta_r\Delta_r^{-1}\check{\Phi}.$$

Obviously, for $t=t'$ the diagonal elements of $\rho_{12}^{(n)}$ are positive. If we go from the variable ζ to the arguments $\mathbf{N}$ with the aid of the function Φ^0, then it turns out that

$$\rho_{12}^{(n)}(\mathbf{N},t,\mathbf{N},t)>0.$$

That the diagonal elements of $\rho_{12}^{(c)}$ are positive definite is almost obvious. For the free electromagnetic field we can write

$$\rho_{12}^0=\chi^0(\chi^0)^+.$$

Then, according to Eq. (13),

$$\rho_{12}^{(c)}=\chi\chi^+, \quad \chi=(1+\Delta_r\hat{\mathscr{P}}_r)\chi^0.$$

Now it is clear that

$$\rho_{12}^{(c)}(\mathbf{N},t,\mathbf{N},t)=\chi(\mathbf{N},t)\chi^+(\mathbf{N},t)>0.$$

If

$$\hat{\mathbf{s}}=\frac{1}{V}\sum_{\mathbf{k}\lambda}\mathbf{k}\hat{\alpha}^+_{\mathbf{k}\lambda}\hat{\alpha}_{\mathbf{k}\lambda} \tag{24}$$

is the operator expression for the Poynting vector and

$$\langle\mathbf{s}\rangle=\mathrm{Tr}(\hat{\mathbf{s}}\rho_{12})=\sum_N\hat{\mathbf{s}}\rho_{12}(\mathbf{N},t,\mathbf{N},t),$$

then from Eqs. (12)–(14) and (24), we find

$$\langle\mathbf{s}\rangle=\langle\mathbf{s}\rangle^{(c)}+\langle\mathbf{s}\rangle^{(n)},$$

where

$$\langle\mathbf{s}\rangle^{(c)}=\sum_{\mathbf{k}\lambda}(\mathbf{s})^{(c)}_{\mathbf{k}\lambda}, \quad \langle\mathbf{s}\rangle^{(c)}_{\mathbf{k}\lambda}=\sum_{\mathbf{N}}\frac{N_{\mathbf{k}\lambda}}{V}\mathbf{k}\rho_{12}^{(c)}(\mathbf{N},t,\mathbf{N},t), \tag{25a}$$

$$\langle\mathbf{s}\rangle^{(n)}=\sum_{\mathbf{k}\lambda}\langle\mathbf{s}\rangle^{(n)}_{\mathbf{k}\lambda}, \quad \langle\mathbf{s}\rangle^{(n)}_{\mathbf{k}\lambda}=\sum_{\mathbf{N}}\frac{N_{\mathbf{k}\lambda}}{V}\mathbf{k}\rho_{12}^{(n)}(\mathbf{N},t,\mathbf{N},t). \tag{25b}$$

The components of the Poynting vector corresponding to the coherent and incoherent scattering channels add vectorially. Here there are no interference phenomena. Of greatest interest is Eq. (25b), which indicates that the component of the Poynting vector corresponding to the incoherent channel can be written as a sum of components over different modes, each of which, in turn, can be expanded as a sum of positive definite components corresponding to different configurations $\mathbf{N}$ of the scattered field. Here, also, there are no interference terms. Thus, if we limit the sum over $\mathbf{N}$ in Eq. (25b) to only the chosen configurations, we obtain a lower bound on the absolute value of the vector $\langle\mathbf{s}\rangle^{(n)}_{\mathbf{k}\lambda}$.

All these remarks apply fully to the reflection coefficient, R as well. Below, in studying the incoherent scattering channel we limit ourselves to reflection processes that result in the production of two photons in the reflected wave mode $(\mathbf{k},\lambda)$. Including only these processes makes it possible, on the one hand, to obtain an estimated lower bound on the reflection coefficient $\mathbf{R}$ and, on the other, to reveal the characteristic behavior of the incoherent scattering channel. Specifically, we shall have in mind processes involving the scattering of photons by unexcited atoms, and the subsequent stimulated emission of atoms excited by the scattered photons. Their contribution to R sets in only with the sixth order of perturbation theory. Then, in accordance with the discus-

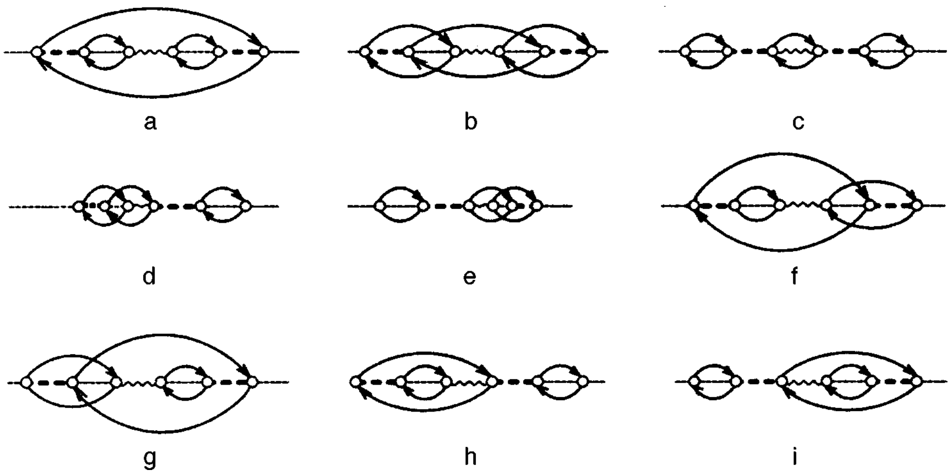


FIG. 3. Diagrams of processes that put two photons in the reflection mode.

sion of Sec. 3, here we should expect the appearance of singularities requiring the summation of an infinite subsequence of Feynman diagrams.

Now it is clear that in this sort of problem, the contribution of infinite subsequences of Feynman diagrams whose first term shows up only in the sixth order of perturbations may end up being significant. The diagrams of Fig. 3 correspond to this process in the Green function Γ-technique. The arrows have been omitted in the retarded and advanced Green functions, as well as in ρ^0_{12}. The sum of the expressions corresponding to these diagrams yields the following formula for the reflection coefficient:

$$R^{\lambda\lambda_0(n)}_{\mathbf{kk}_0}=\frac{\pi^6(2j_m+1)^3\gamma_r^3 n_m\delta_{\lambda\lambda_0}}{L_z\omega^3_{m\mu}k_0^2k_{0z}^3}\sin^2(k_{0z}L)\,\delta(k_{0x},k_x)\times\delta(k_{0y},k_y)\,\delta_\gamma(2k-k_0-\omega_{m\mu})\,\delta^2_\Gamma(k_0-k)\times\left|\frac{n_\mu}{k-\omega_{m\mu}+i\gamma/2}\right|^2,\quad n_m\ll n_\mu. \tag{26}$$

For the polarization operator $\hat{\mathscr{P}}^{(n)}_{12}$ we have used the diagonal approximation, which corresponds to an infinitely thick scattering layer. Equation (26) corresponds to the sixth order proper of perturbation theory, with δ_Γ taken to mean the standard Dirac δ-function. The square of this function characterizes precisely the singularity mentioned above. In order to eliminate the singularity, we sum an infinite subsequence of Feynman diagrams, which is denoted in Fig. 3 by thick dashed lines. A thick line denotes the complete Green function of a scattered $(\mathbf{k},\lambda)$ photon produced by the stimulated emission from an excited atom (Fig. 3a). Also possible is a process of spontaneous emission into the $(\mathbf{k},\lambda)$ mode, which then induces stimulated scattering of the initial field $(\mathbf{k}_0,\lambda_0)$ (Fig. 3b–3e). These processes can interfere (Fig. 3f–3i). In all cases, the Green function Δ^0_r of the photon which provokes the stimulated processes is replaced by Δ_r. An expression valid for an infinite medium is used for Δ_r. In energy space this substitution has the form

$$\Delta^0_r\to\Delta_r=\frac{1}{E-k-\hat{\mathscr{P}}_r(E)},$$

where $\hat{\mathscr{P}}_r(E)$ is the polarization operator (17) in the **N**-representation when a single photon is present in the $(\mathbf{k},\lambda)$ mode. Accordingly, δ_Γ is understood to be the expression

$$\delta_r(k_0-k)=-\frac{1}{2\pi i}\left[\frac{1}{k_0-k-\hat{\mathscr{P}}_r(k_0)}-\frac{1}{k_0-k-\hat{\mathscr{P}}_a(k_0)}\right].$$

For single-mode occupation with $n_m\ll n_\mu$, according to Eqs. (17) and (18) we have

$$\hat{\mathscr{P}}_r(k_0)=c_r^{\mathbf{k}\lambda}(k_0)=\frac{\pi(2j_m+1)\gamma_r}{2\omega^2_{m\mu}}\,\frac{n_\mu}{k_0-\omega_{m\mu}+i\gamma/2}.$$

As Eq. (26) implies, in the reflected ray the components k_x and k_y coincide with the corresponding components of the scattered flux, as dictated by the geometry of the problem. The component k_z may differ somewhat from $-k_{0z}$. The difference depends on the lifetime of the excited atom (γ^{-1}) and lifetime of the one-photon state in the medium $(\mathrm{Im}\,\hat{\mathscr{P}}_r)^{-1}$. Because of the resulting nonuniqueness of k_z, the reflected wave acquires a distribution in frequency and angle.

We are interested in the relative intensities of the coherent and incoherent channels. For the coherent channel in the case of resonance frequencies, according to Eq. (19), we have

$$R^{(c)}=\delta_{\lambda\lambda_0}\,\frac{\pi^2\gamma_r^2(2j_m+1)(n_\mu+n_m)^2}{k^4_{0z}\omega^2_{m\mu}\gamma^2}\sin^2(k_{0z}L).$$

A number of situations are possible for the incoherent channel. Let

$$|\mathscr{P}_r(k)|^2_{\max}=\left|\frac{\pi(2j_m+1)}{\omega^2_{m\mu}}\,\frac{\gamma_r}{\gamma}\,n_\mu\right|^2\ll\gamma^2, \tag{27}$$

i.e.,

$$n_\mu\bar{\lambda}^3\,\frac{k_0\gamma_r}{\gamma^2}\ll 1.$$

According to Eq. (26), for $n_m \ll n_\mu$ and $k_{0z} \gg k_{0x}, k_{0y}$ we now have

$$\sum_{\mathbf{k}} R_{\mathbf{kk}_0}^{\lambda\lambda_0(n)} = \frac{4n_m}{n_\mu} R^{(c)}. \tag{28}$$

Thus, for $n_m \ll n_\mu$ the contribution of the incoherent channel can be neglected. When $n_m \sim n_\mu$, it must be taken into account. And if perturbation terms beginning with the sixth order contribute to $R^{(n)}$, their overall result will be proportional to e^4, owing to the above-mentioned singularities after summation of the diagrams. Even though with increasing n_m the contribution of the incoherent channel becomes comparable to that of the coherent channel according to Eq. (28), under real thermal excitation conditions this is scarcely possible. In fact, when the inequality (27) is satisfied, light that is coherently reflected from a surface at a specular angle will be lost among the diffusely reflected radiation. The specular component becomes significant when[29]

$$n_\mu \lambdabar^3 \frac{\gamma_r}{\rho} \sim 1.$$

Inclusion of the Doppler effect, which facilitates the inequality (27), can hardly change the situation, at least in the visible region of the spectrum. As n_μ increases, the inequality (27) is violated. With the experimentally easily attainable inequality opposite to Eq. (27), we have

$$\sum_{\mathbf{k}} R_{\mathbf{kk}_0}^{\lambda\lambda_0(n)} \propto \frac{n_m}{(2j_m+1)n_\mu} R^{(c)} \frac{\gamma^2}{k_0 \gamma_r n_\mu \lambdabar^3}.$$

Under these conditions the integrated relative contribution of two-photon processes is negligible.

5. CONCLUSION

Since the reflection coefficient in the coherent channel, $R^{(c)} \propto (n_\mu + n_m)^2$, depends only on the total concentration of scatterers, but not on n_μ and n_m individually, we have $dR^{(c)}/dn_m = 0$. Allowance for the incoherent channel at small n_m leads to the inequality

$$\frac{dR}{dn_m} = \frac{d}{dn_m}(R^{(c)} + R^{(n)}) = \frac{dR^{(n)}}{dn_m} \geqslant 0, \tag{29}$$

i.e., to a rise in the total selective reflection coefficient as the concentration of excited scatterers increases. As this analysis has shown, under real conditions the equality in Eq. (29) is effectively realized and the total reflection coefficient loses its linear dependence on n_m. In addition, in the semiclassical theory of radiation, we have

$$\frac{dR}{dn_m} < 0, \quad \text{for } n_m < n_\mu,$$

since $R \propto (n_\mu - n_m)^2$ in this theory.

The predictions of semiclassical radiation theory and rigorous quantum electrodynamics are compared in Fig. 4. There is a qualitative disagreement between them, which does not favor the semiclassical theory. Experimental studies[15–18] thus far have had the goal of attaining $R > 1$, and were conducted on media with a population inversion, $n_m > n_\mu$. The simpler thermal excitation region that we have examined lay outside the field of view of the experimenters. But even here, in highly inhomogeneous media, the stimulated emission behaves in a nontrivial fashion. As n_m increases, the discrepancy between the theories becomes greater, and in the population inversion region the calculations should be exclusively quantum electrodynamical in character.

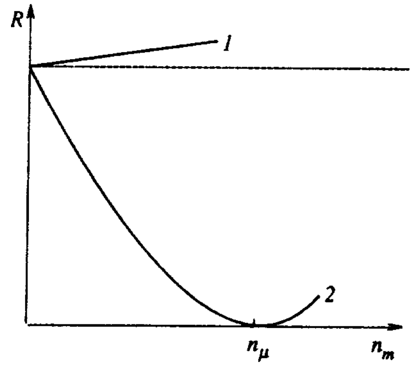


FIG. 4. The selective reflection coefficient as a function of the degree of excitation of a medium: (*1*) quantum theory of reflection; (*2*) semiclassical theory of reflection.

[1] C. Kenty, Phys. Rev. **42**, 823 (1932); **43**, 181 (1933).
[2] L. M. Biberman, Zh. Éksp. Teor. Fiz. **17**, 416 (1947).
[3] T. Holstein, Phys. Rev. **72**, 1212 (1947).
[4] I. M. Petrov, Yu. A. Matyugin, and V. P. Chebotarev, JETP Lett. **10**, 187 (1969); Opt. i Spektrosk. **30**, 356 (1970).
[5] M. I. D'yakonov and V. I. Perel', Zh. Éksp. Teor. Fiz. **58**, 1090 (1970) [Sov. Phys. JETP **31**, 585 (1970)].
[6] I. M. Petrov, Yu. A. Matyugin, S. G. Rautian, and V. P. Chebotarev, Zh. Éksp. Teor. Fiz. **58**, 1243 (1970) [Sov. Phys. JETP **31**, 668 (1970)].
[7] A. P. Kazantsev, V. S. Smirnov, and V. P. Yakovlev, Zh. Éksp. Teor. Fiz. **82**, 1738 (1982) [Sov. Phys. JETP **55**, 1004 (1982)].
[8] A. P. Kazantsev, O. G. Malikyan, and V. P. Yakovlev, Zh. Éksp. Teor. Fiz. **89**, 450 (1985) [Sov. Phys. JETP **62**, 255 (1985)].
[9] A. A. Panteleev, V. A. Roslyakov, and A. N. Starostin, Zh. Éksp. Teor. Fiz. **97**, 1777 (1990) [Sov. Phys. JETP **70**, 1003 (1990)].
[10] Yu. K. Zemtsov, A. Yu. Sechin, and A. N. Starostin, Zh. Éksp. Teor. Fiz. **110**, 1654 (1996) [JETP **83**, 909 (1996)].
[11] L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1964)].
[12] B. A. Veklenko, Izv. Vuzov, Fizika, No. 9, 71 (1983).
[13] B. A. Veklenko and G. B. Tkachuk, Izv. Vuzov, Fizika, No. 2, 89 (1987).
[14] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Physics* [in Russian], Gostekhizdat, Moscow (1962).
[15] Ch. I. Koester, IEEE **2**(9), 580 (1966).
[16] S. A. Lebedev, V. A. Kizel', and B. Ya. Kogan, Kvant. Élektron. **3**, 2446 (1976) [Sov. J. Quantum Electron. **6**, 1332 (1976)].
[17] B. B. Boĭko and N. N. Uvarova, Kvant. Élektron. **8**, 2506 (1981) [Sov. J. Quantum Electron. **11**, 1530 (1981)].
[18] B. B. Boĭko and N. S. Petrov, *Reflection of Light from Amplifying and Nonlinear Media* [in Russian], Nauka i Tekhnika, Minsk (1988).
[19] G. N. Romanov and S. S. Shakhidzhanov, JETP Lett. **16**, 210 (1972).
[20] A. A. Kolokolov, JETP Lett. **21**, 312 (1975).
[21] G. N. Vinokurov and V. I. Zhulin, Kvant. Élektron. **9**, 553 (1982) [Sov. J. Quantum Electron. **12**, 329 (1982)].
[22] B. B. Boĭko, I. Z. Dzhilavdari, and N. S. Petrov, Zh. Prikl. Spektrosk. **25**, 148 (1976).
[23] A. B. Zimin, Zh. Prikl. Spektrosk. **40**, 1005 (1984).
[24] B. A. Veklenko, Zh. Éksp. Teor. Fiz. **96**, 457 (1989) [Sov. Phys. JETP **69**, 258 (1989)].
[25] R. Shoemaker, in *Laser and Coherent Spectroscopy* [Russian translation], J. Steinfeld (ed.), Mir, Moscow (1982).

[26] W. Heitler, *The Quantum Theory of Radiation* , Oxford Univ. Press, New York (1954).
[27] A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics* [in Russian], Nauka, Moscow (1969).
[28] V. P. Silin and V. Ya. Faĭnberg, Usp. Fiz. Nauk **61**, 568 (1955).
[29] L. M. Biberman, Pure Appl. Chem. **13**, 393 (1966).

Translated by D. H. McNeill

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Two-electron excitations in helium-like ions by inelastic photon scattering

M. Ya. Amus'ya, A. I. Mikhaĭlov,*) and I. A. Mikhaĭlov

B. P. Konstantinov St. Petersburg Institute of Nuclear Physics, Russian Academy of Sciences, 188350 Gatchina, Leningrad Region, Russia

(Submitted 28 April 1997)
Zh. Éksp. Teor. Fiz. **113**, 539–549 (February 1998)

Inelastic photon scattering by helium atoms and helium-like ions with simultaneous excitation of the two-electron transition $1s^2 \rightarrow 2s^2$ is examined in the nonrelativistic energy range $I \ll \omega \ll m$ (I is the ionization potential, ω is the photon energy, m is the electron mass, and $\hbar = c = 1$). The electrons are assumed to be moving in the Coulomb field of the nucleus, and the electron–electron interaction is taken into account in the lowest perturbation order. The differential and total cross sections of the process and the autoionization width of the $2s^2$ energy level are calculated. The numerical value of the autoionization width is found to agree with the results of the more rigorous calculations of other researchers. © *1998 American Institute of Physics.* [S1063-7761(98)01102-0]

1. INTRODUCTION

The two-electron transitions in an atom induced in collisions with photons or charged particles have been intensively studied for about 30 years.[1–8] The importance of such studies is related to the possibility of effectively extracting information about the role of electron correlations in the atom. When there is a collision with a photon, two-electron transitions are determined solely by the electron–electron interaction. Among such transitions are double ionization, ionization with excitation, and double excitation of an atom, which can be observed both in the scattering of a photon and in photon absorption. Double excitation in photoabsorption can occur only at resonance frequencies of the incident radiation, i.e., at low photon energies ($\omega \sim I$), which we do not consider here. At photon energies $\omega < \eta$ (here η is the average momentum of the atomic electron) the dominant processes are those with photoabsorption, while at $\omega > \eta$ the dominant processes are those with photon scattering.[6,7]

The simplest two-electron systems are the helium atom and helium-like ions, and it is largely this simplicity that explains the interest in them. The processes that have been studied most thoroughly are double photoionization, the removal of two atomic electrons in the absorption of a single photon,[1–4,9–12] and the double Compton effect—the double ionization in the scattering of a photon by atomic electrons.[13–19] Ionization with excitation in photoabsorption was studied in Refs. 4, 10, 20 and 21, and the analogous two-electron transition in Compton scattering was studied in Refs. 15 and 21. In recent years there has also been an upsurge of interest in the problem of double excitation of atoms by fast charged particles. Among the various papers devoted to this problem we mention Refs. 5 and 22–25.

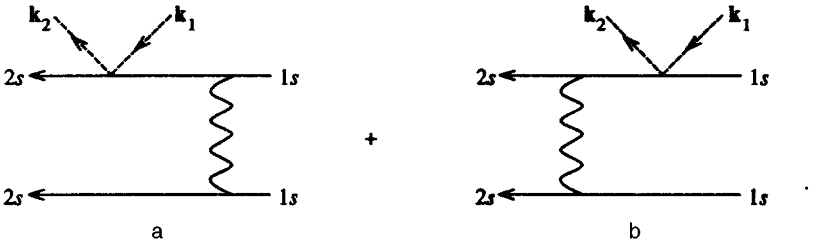
This paper examines for the first time the process of double excitation of an atom by high-frequency photons ($\omega \gg I$). We study the two-electron transition $1s^2 \rightarrow 2s^2$ in the helium atom and in helium-like ions in the inelastic scattering of x-ray photons ($I \ll \omega \ll m$). The electron–electron interaction is taken into account in the first perturbation order. In the zeroth approximation, for electrons we take the Coulomb wave functions and the Coulomb Green's function. We find the angular distribution of the scattered photons and the total cross section of formation of the autoionization $2s^2(^1S)$ state of the helium atom. Using the same model, we calculate the width of the $2s^2$ level for Auger decay, or the autoionization width. We find that there is good agreement between our results and those found in the literature for this quantity,[26,27] which justifies our approach.

It turns out that in the adopted approximation there is no forward scattering of photons. The total cross section of the process exhibits a sharp peak at $\omega \sim \eta$. The peak value of the cross section is $\sigma_{\max} \approx 4 \times 10^{-28} Z^{-2}$ cm^2 (Z is the charge of the nucleus). It is interesting to compare the cross section of double and single excitations of the helium atom in the scattering of photons of equal energies. The formula for the cross section of the one-electron transition $1s \rightarrow 2s$ was obtained in Ref. 28. A calculation by this formula for $\omega = \eta$ yields[1)] $\sigma(1s^2 \rightarrow 1s2s) \approx 2 \times 10^{-26}$ cm^2, which is 200 times the cross section of the corresponding two-electron transition for $Z = 2$. This means that the electron–electron interaction in helium is relatively weak.

One must bear in mind that the final products of the decay of the autoionization $2s^2$ state are a singly charged ion and an electron. Hence the process under investigation provides a contribution (albeit insignificant) to the cross section of formation of singly charged ions He$^+$ by photon impact.

2. DERIVING THE EXPRESSION FOR THE PROCESS AMPLITUDE

Let us consider electrons whose energies ω lie in the range $I \ll \omega \ll m$. As shown in Ref. 16, the A^2-term in the electron–photon interaction operator plays the leading role in one- and two-electron ionization in Compton scattering. The same is true of ionization with excitation when a photon is scattered by an atom.[21] Proceeding in the same way as in Ref. 16, we can show that the A^2-term is also the leading one in the above energy range for double excitation by photon scattering (DEPS). By way of an example, we study the two-electron transition from the ground state to the nearest excited state of the helium atom or the helium-like ion, the $2s^2$ state. Bearing in mind what was said earlier, we depict the DEPS amplitude by the following Feynman diagrams:



(1)

Here the dashed lines with arrows denote photons with momenta $\mathbf{k}_1$ and $\mathbf{k}_2$ and the solid lines, electrons. A wavy line denotes the electron–electron interaction. A vertex with two photon lines corresponds to the operator $m^{-1}\mathbf{A}_2^*\mathbf{A}_1$ (the A^2-term). In the momentum representation, which we use in our calculations, this operator is represented by a factor $(4\pi\alpha/\sqrt{4\omega_1\omega_2})\cdot(\mathbf{e}_2^*\cdot\mathbf{e}_1)/m$, where $\alpha\cong 1/137$ is the fine structure constant, and $\mathbf{e}_1,\omega_1$ $(\mathbf{e}_2,\omega_2)$ are the polarization vector and energy of a photon. Below, this factor and the factor $4\pi\alpha$ from the electron–electron interaction will be extracted from the amplitude and introduced explicitly into the formulas for the cross section. The diagram (1a) takes into account the electron–electron interaction in the initial state and the diagram (1b), in the final state. Since the photons interact with each electron separately, the total amplitude must have two more diagrams with a photon vertex on the lower electron line. The contribution of such diagrams is equal to the contribution of (1). Hence the total amplitude is twice the amplitude (1).

Earlier (see Refs. 16 and 21) we found that in processes involving ionization, the diagram allowing for the electron–electron interaction in the final state yields a small contribution compared to the diagram that allows for this interaction in the initial state. Below we will show that in DEPS processes both diagrams are important, the reason for this being that in the final state both electrons have an energy of order I.

Thus, we can write the amplitude of the process in the following form:

$$M=2(M_a+M_b), \tag{2}$$

$$M_{a(b)}=\int\frac{d\mathbf{f}}{(2\pi)^3}\,\frac{1}{f^2}F(\mathbf{f})F_{a(b)}(\mathbf{f}), \tag{3}$$

$$F(\mathbf{f})=\int\frac{d\mathbf{f}'}{(2\pi)^3}\langle 2s|\mathbf{f}'\rangle\langle\mathbf{f}'+\mathbf{f}|1s\rangle, \tag{4}$$

$$F_a(\mathbf{f})=\int\frac{d\mathbf{f}_2\,d\mathbf{f}_1}{(2\pi)^6}\langle 2s|\mathbf{k}+\mathbf{f}_2\rangle\langle\mathbf{f}_2|G_a|\mathbf{f}_1\rangle\langle\mathbf{f}_1-\mathbf{f}|1s\rangle, \tag{5}$$

$$F_b(\mathbf{f})=\int\frac{d\mathbf{f}_2\,d\mathbf{f}_1}{(2\pi)^6}\langle 2s|\mathbf{f}+\mathbf{f}_2\rangle\langle\mathbf{f}_2|G_b|\mathbf{f}_1\rangle\langle\mathbf{f}_1-\mathbf{k}|1s\rangle, \tag{6}$$

with $\mathbf{k}=\mathbf{k}_1-\mathbf{k}_2$ the momentum transferred to the atom, and $G_a\equiv G_c(E_a)$ the Coulomb Green's function for the electron in diagram (1a).

The function G_b is defined similarly. By virtue of energy conservation,

$$E_a=2E_{1s}-E_{2s}=-\tfrac{7}{4}I,\quad E_b=2E_{2s}-E_{1s}=\tfrac{1}{2}I. \tag{7}$$

To calculate F and $F_{a(b)}$ we need the Coulomb wave functions in the momentum representation:

$$\langle\mathbf{f}'-\mathbf{f}|1s\rangle=N_1\left(-\frac{\partial}{\partial\eta_1}\right)\langle\mathbf{f}'|V_{i\eta_1}|\mathbf{f}\rangle,$$

$$N_1^2=\frac{\eta_1^3}{\pi},\quad \eta_1=\eta=m\alpha Z,$$

$$\langle\mathbf{f}'-\mathbf{f}|2s\rangle=N_2\Gamma_{\eta_2}\left(-\frac{\partial}{\partial\eta_2}\right)\langle\mathbf{f}'|V_{i\eta_2}|\mathbf{f}\rangle,$$

$$N_2^2=\frac{\eta_2^3}{\pi},\quad \eta_2=\frac{\eta}{2},$$

$$\Gamma_{\eta_2}=1+\frac{\eta}{2}\frac{\partial}{\partial\eta_2},\quad \langle\mathbf{f}'|V_\lambda|\mathbf{f}\rangle=\frac{4\pi}{(\mathbf{f}'-\mathbf{f})^2-\lambda^2}. \tag{8}$$

Now we can insert (8) into (4)–(6) and perform several integrations. The result is

$$M_a=\hat D\left(-\frac{\partial}{\partial\eta_2}\right)\frac{1}{\lambda^2}\langle\mathbf{k}|V_{i\eta_2}G_aV_{i\mu}|0\rangle,\quad \mu=\eta_1+\lambda, \tag{9}$$

$$M_b=\hat D\left(-\frac{\partial}{\partial\eta_1}\right)\frac{1}{\lambda^2}\langle\mathbf{k}|V_{i\eta_1}G_bV_{i\nu}|0\rangle,\quad \nu=\eta_2+\lambda, \tag{10}$$

$$\hat D=N^2\Gamma_{\eta_2}\Gamma_\lambda\frac{\partial}{\partial\lambda},\quad \Gamma_\lambda=1+\frac{\eta}{2}\frac{\partial}{\partial\lambda},$$

$$N=N_1N_2,\quad \lambda=\frac{3}{2}\eta. \tag{11}$$

The matrix elements in (9) and (10) can be represented as simple integrals.[29] Here is the integral representation of the matrix element in (10):

TABLE I.

θ	$-\alpha Z M_a \times 10^3$	αZ Re $M_b \times 10^3$	$-\alpha Z$ Im $M_b \times 10^3$
0	2.3287	2.3287	0.0000
18	2.5964	2.4123	0.1395
36	3.1085	2.5346	0.5711
54	3.3998	2.4102	1.1697
72	3.3460	1.9713	1.6092
90	3.0806	1.4548	1.7408
108	2.7602	1.0476	1.6735
126	2.4781	0.7813	1.5449
144	2.2700	0.6242	1.4295
162	2.1459	0.5434	1.3553
180	2.1049	0.5185	1.3301

$$J_b = \langle \mathbf{k} | V_{i\eta_1} G_b V_{i\nu} | 0 \rangle = 16\pi i p m \int_0^1 \frac{t^{-i\xi} dt}{\alpha - 2\beta t + \gamma t^2}, \tag{12}$$

$$p = p_b = \sqrt{2mE_b} = \frac{\eta}{\sqrt{2}}, \quad \xi = \frac{\eta}{p} = \sqrt{2}, \tag{13}$$

$$\alpha = \alpha_1 \alpha_2, \quad \alpha_1 = (\nu - ip)^2, \quad \alpha_2 = k^2 + (\eta_1 - ip)^2,$$

$$\gamma = \gamma_1 \gamma_2, \quad \gamma_1 = (\nu + ip)^2, \quad \gamma_2 = k^2 + (\eta_1 + ip)^2,$$

$$\beta_1 = \beta_1 \beta_2, \quad \beta_1 = \nu^2 + p^2, \quad \beta_2 = k^2 + \eta_1^2 + p^2. \tag{14}$$

The integral representation for $J_a = \langle \mathbf{k} | V_{i\eta_2} G_a V_{i\mu} | 0 \rangle$ can be obtained from (12) by replacing η_1 with η_2, ν with μ, and p_b with p_a. Since $p_a = \sqrt{2mE_a} = i\sqrt{2m|E_a|}$, the last replacement is equivalent to $p \rightarrow iq$ and $i\xi \rightarrow \zeta = \eta/q$, where because of (7) $q = \sqrt{2m|E_a|} = \eta\sqrt{7}/2$. Clearly, after these transformations have been carried out, the integral J_a becomes real, and so does the entire amplitude M_a. Note that the amplitude M_b is complex-valued.

The integral (12) can be expressed in terms of hypergeometric functions, but the resulting calculations of four derivatives in (9) and (10) lead to cumbersome expressions. We believe that there is a more rational path of reasoning. First we calculated all the derivatives and then integrated numerically. But still the expression obtained after calculating the derivatives is too cumbersome to write it here, so we mention only several characteristic features of this expression. The amplitude M is dimensionless. If the photon energy is expressed in units of η, the dependence of the amplitude on Z is very simple: $M \propto Z^{-1}$. The angular dependence enters into the amplitude through the square of the momentum transfer:

$$k^2 = \omega_1^2 + \omega_2^2 - 2\omega_1\omega_2 \cos\theta. \tag{15}$$

Since

$$\omega_1 - \omega_2 = 2(E_{2s} - E_{1s}) = \tfrac{3}{2} I = \tfrac{3}{4} m\alpha^2 Z^2,$$

we can assume that $\omega_2 = \omega_1$ and put

$$k^2 = 2\omega_1^2 (1 - \cos\theta).$$

The dependence of the amplitudes M_a and M_b on the scattering angle θ calculated by (9) and (10) for the photon energy $\omega_1 = \eta$ is listed in Table I.

Here are two facts that follow from Table I:

1) The amplitudes M_a and Re M_b have close values but opposite signs, which leads to partial cancellation of their contribution.

2) At $\theta = 0$ (forward scattering) $M_a = -M_b$, which means that the amplitudes (1a) and (1b) balance each other perfectly.

Of course, it is advisable to corroborate the numerical results analytically, which we do in the next section.

3. FORWARD SCATTERING AMPLITUDE

It turns out that simple expressions can be derived for the amplitudes M_a and M_b at $\theta = 0$. These expressions can be used to check the accuracy of the numerical calculations. Let us go back to the general formulas (9) and (10) and see how they are altered when $k = 0$ (at $\theta = 0$ the square of the momentum transfer, k^2, is approximately $(\alpha Z)^2 \eta^2$, and on the basis of (12) and (14) we can put $k = 0$). Using the expressions (8) for the $1s$- and $2s$-wave functions and the expansion of the Coulomb Green's function in the eigenfunctions of the Schrödinger equation,

$$G_c(E) = \sum_n \frac{|\psi_n\rangle\langle\psi_n|}{E - E_n}, \tag{16}$$

we obtain

$$M_a = \hat{A} N_2 \Gamma_{\eta_2} \left(-\frac{\partial}{\partial\eta_2} \right) \langle 0 | V_{i\eta_2} G_a V_{i\mu} | 0 \rangle$$

$$= \hat{A} \langle 2s | G_a V_{i\mu} | 0 \rangle = \frac{\hat{A} \langle 2s | V_{i\mu} | 0 \rangle}{E_a - E_{2s}},$$

$$\hat{A} = N_1^2 N_2 \Gamma_\lambda \frac{\partial}{\partial\lambda} \frac{1}{\lambda^2}, \tag{17}$$

$$M_b = \hat{B} N_1 \left(-\frac{\partial}{\partial\eta_1} \right) \langle 0 | V_{i\eta_1} G_b V_{i\nu} | 0 \rangle$$

$$= \hat{B} \langle 1s | G_b V_{i\nu} | 0 \rangle = \frac{\hat{B} \langle 1s | V_{i\nu} | 0 \rangle}{E_b - E_{1s}},$$

$$\hat{B} = N_1 N_2^2 \Gamma_{\eta_2} \Gamma_\lambda \frac{\partial}{\partial\lambda} \frac{1}{\lambda^2}. \tag{18}$$

The energies E_a and E_b are defined in (7).

We introduce the notation $\omega = \omega_1 - \omega_2$. Then

$$E_a - E_{2s} = 2(E_{1s} - E_{2s}) = -\omega,$$

$$E_b - E_{1s} = 2(E_{2s} - E_{1s}) = \omega. \tag{19}$$

Applying the operators $\hat{A}$ and $\hat{B}$ to the matrix elements to the right of these operators, we arrive at the following equality:

$$\hat{A} \langle 2s | V_{i\mu} | 0 \rangle = \hat{B} \langle 1s | V_{i\nu} | 0 \rangle = \frac{\pi}{3\lambda^5} N^2. \tag{20}$$

Inserting (19) and (20) into (17) and (18) and allowing for the relationships

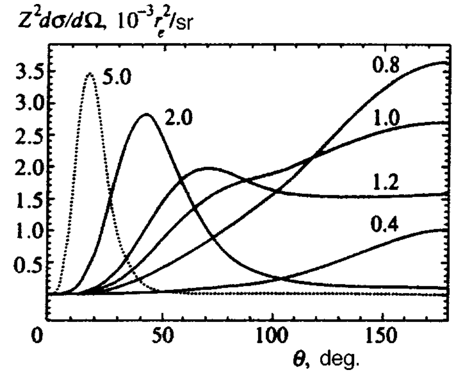


FIG. 1. Angular distributions of scattered photons for different values of the energy ω_1 of the incident photon. The curves are labeled with values of ω_1 in units of η.

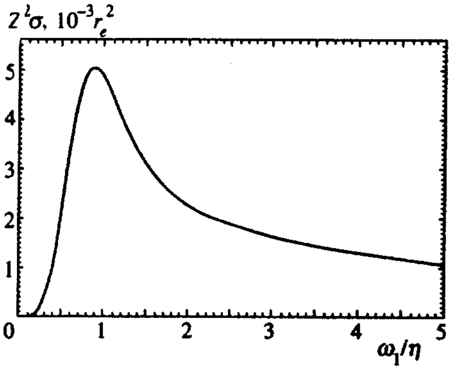


FIG. 2. Total cross section of double excitation of the $2s^2$ state of a helium-like ion as a function of incident photon energy.

$$\omega=\frac{3}{4}\alpha Z\eta, \quad \lambda=\frac{3}{2}\eta, \quad N^2=\frac{\eta^6}{8\pi^2}, \tag{21}$$

we find that

$$M_b(0)=-M_a(0)=\frac{1}{\alpha Z}\frac{1}{8\pi}\left(\frac{2}{3}\right)^7=\frac{2.3287\times10^{-3}}{\alpha Z}, \tag{22}$$

which coincides with the result listed in Table I ($M_{a(b)}(0)$ denotes the value of $M_{a(b)}$ at $\theta=0$).

Since the amplitude M of the process is $2(M_a+M_b)$, there is no forward scattering of photons.

4. ANGULAR DISTRIBUTION OF PHOTONS AND THE TOTAL CROSS SECTION

The differential cross section of the process is

$$d\sigma=(4\pi\alpha)^2(4\pi r_e)^2\frac{|\mathbf{e}_2^*\cdot\mathbf{e}_1|^2}{4\omega_1\omega_2}|M|^2\frac{d^3k_2}{(2\pi)^3}$$
$$\times 2\pi\delta(2E_{2s}+\omega_2-\omega_1-2E_{1s}), \tag{23}$$

where $r_e=\alpha/m$ is the classical electron radius ($r_e^2=7.95\times10^{-26}$ cm^2). Summing and averaging over photon polarizations and integrating with respect to ω_2, we arrive at the following expression for the angular distribution of the scattered photons:

$$\frac{d\sigma}{d\Omega}=r_e^2\frac{32\pi^2}{Z^2}(1+\cos^2\theta)|\alpha Z(M_a+M_b)|^2, \tag{24}$$

where θ is the angle between the photon momenta $\mathbf{k}_1$ and $\mathbf{k}_2$. As (22) clearly shows, the product of αZ and the amplitude ceases to depend on Z if the photon energies are expressed in units of $\eta=m\alpha Z$. From the distribution (24) over the photon scattering angle we can proceed to the distribution over the momentum transfer if we replace $\cos\theta$ by $(1-k^2/2\omega_1^2)$ and $d\Omega$ by $\pi d(k^2/\omega_1^2)$.

The total cross section can be written as

$$\sigma=r_e^2\frac{(4\pi)^3}{Z^2}\int_{-1}^{+1}(1+x^2)|\alpha Z(M_a+M_b)|^2dx, \tag{25}$$

with $x=\cos\theta$. The angular distribution of the scattered photons calculated by (24) is depicted in Fig. 1 for various values of ω_1/η (the labels on the curves indicate these values). What is interesting here is that for $\omega_1<\eta$ the angular distributions have a broad maximum at $\theta=180°$. For values of ω_1 slightly larger than η, the maximum shifts to angles smaller than 90°. If the energy is increased still further, the maximum in the distribution shifts to smaller angles and narrows.

The dependence of the total cross section on the incident-photon energy is shown in Fig. 2. At $\omega_1\simeq0.9\eta$ the cross section has a peak value of $\sigma\approx4\times10^{-28}Z^{-2}$ cm^2. Here we also list the values of the cross section at two limits:

$$\omega_1=0.1\eta: \quad \sigma=4.55\times10^{-31}Z^{-2}\ \text{cm}^2,$$

$$\omega_1=10\eta: \quad \sigma=4.26\times10^{-29}Z^{-2}\ \text{cm}^2$$

(such values of ω_1 land in the range $I\ll\omega_1\ll m$ only if Z is small). The drop in cross section with decreasing $\omega_1<\eta$ can be explained by the fact that k^2 becomes small for all angles, and because of this the contributions of the diagrams (1a) and (1b) are almost balanced. The cross section also drops as $\omega_1>\eta$ increases, which is due to a concentration of the process near $k\sim\eta$ or $\theta\sim\eta/\omega_1$ (see Eqs. (12)–(15).

Qualitatively, such behavior of the cross sections (differential and total) can be understood if we turn to Compton scattering by a free electron, where the energy Δ lost by the photon is determined by the scattering angle θ (see Ref. 30):

$$\Delta=\frac{\omega_1^2}{m}(1-\cos\theta).$$

The rigorous relationship between these kinematic quantities is due to energy and momentum conservation. In scattering by a bound electron, the correlation between the photon and electron variables is determined only by energy conservation, since there is a third body participating in the process, the nucleus, which in view of its large mass is able to absorb any amount of momentum. As a result there emerges the possibility of a process in which a fixed energy is transferred to the atom and the photon can be scattered through any angle. However, the greatest contribution to the cross section is provided by the range of angles closest to free kinematics (the kinematics of scattering by a free electron). For $\omega_1\ll\eta$ (soft photons), even the maximum energy transferred to a free electron ($\Delta_{\max}=2\omega_1^2/m$) is less than the energy $\omega=3I/2$ transferred to the atom in the process. Since scattering through angles $\theta \sim 180°$ exhibits the least deviation from free kinematics, large angles dominate the scattering of soft photons by an atom. The cross section proper is small, since the process occurs in the region that is kinematically inaccessible to scattering by a free electron. For $\omega_1 \gg \eta$ (hard photons) the values $\Delta \sim \omega \sim I$ are possible only at small scattering angles $\theta \sim \eta/\omega_1$. The same angles prevail in inelastic scattering of hard photons by atoms. As the photon energy rises, the small-angle range narrows, and its contribution to the total cross section diminishes.

Photons with energies $\omega_1 \sim \eta$ are the most strongly scattered by an atom. In this case the energy transferred to the atom is close to the energy $\Delta \sim (\eta^2/m)(1-\cos\theta)$ transferred to a free electron over a broad range of scattering angles.

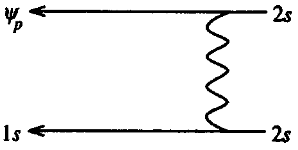
Unfortunately, at present there is nothing to compare our results with. To our knowledge there is not a single work on double excitation in photon scattering. Hence to establish the accuracy of our approach we calculated the rate of Auger decay of the $2s^2$ state of the helium atom under the same assumptions as in calculations of the excitation cross section for this state (i.e., we used the Coulomb wave functions for the electrons and the first perturbation order in the electron–electron interaction) and compared the result with those of other researchers.

5. CALCULATING THE AUTOIONIZATION WIDTH OF THE $2s^2$ LEVEL OF HELIUM

The autoionization width Γ_A and the rate w_A of Auger decay in the system of units in which $\hbar = c = 1$ coincide:

$$\Gamma_A = w_A . \qquad (26)$$

The amplitude M_A of Auger decay can be expressed by the Feynman diagram



(27)

and is given by

$$M_A = \sqrt{2} \int \frac{d\mathbf{f}\, d\mathbf{f}_1\, d\mathbf{f}_2}{(2\pi)^9} \langle \psi_p | \mathbf{f}_1 \rangle \langle \mathbf{f}_1 + \mathbf{f} | 2s \rangle \frac{1}{f^2} \times \langle 1s | \mathbf{f}_2 \rangle \langle \mathbf{f}_2 - \mathbf{f} | 2s \rangle . \qquad (28)$$

The factor $\sqrt{2}$ appears because of the symmetry properties of the wave function of our system.

Here is how it emerges. We use the nonrelativistic approximation, in which the electron–electron interaction is spin-independent and the total wave function is the product of the spatial and spin functions. Because of its orthogonality, the spin part of the wave function does not change in an atomic transition, remains singlet (i.e., antisymmetric), and has no effect on calculations. The spatial part of the wave function in the initial and final states is symmetric:

$$\Psi_i(1,2) = \psi_{2s}(1)\psi_{2s}(2),$$

$$\Psi_f(1,2) = \frac{1}{\sqrt{2}} \{ \psi_p(1)\psi_{1s}(2) + \psi_p(2)\psi_{1s}(1) \}. \qquad (29)$$

The Auger transition amplitude is given by

$$M_A = \langle \Psi_f(1,2) | V(1,2) | \Psi_i(1,2) \rangle . \qquad (30)$$

Since the electron–electron interaction operator $V(1,2)$ is also symmetric under exchange $1 \leftrightarrow 2$ of the electrons, both terms in $\Psi_f(1,2)$ yield the same contribution and we obtain

$$M_A = \sqrt{2} \int \psi_p^*(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2) V(\mathbf{r}_1, \mathbf{r}_2) \times \psi_{2s}(\mathbf{r}_1)\psi_{2s}(\mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2 . \qquad (31)$$

Taking the Fourier transform of (31), we arrive at (28), where $\langle \mathbf{f} | 2s \rangle = \psi_{2s}(\mathbf{f})$, etc.

After the wave functions (8) have been inserted into (28), all integrals can be evaluated, with the result that the amplitude takes the simple form

$$M_A = \sqrt{2}\, \hat{B} \langle \psi_p | V_{i\nu} - V_{i\eta_2} | 0 \rangle . \qquad (32)$$

The operator $\hat{B}$ is defined in (18), $\nu = \eta_2 + \lambda$, $\eta_2 = \eta/2$, and $\lambda = 3\eta/2$. The matrix elements in (32) were calculated in Ref. 21 with ψ_p taken either as a closed expression[31] or as a partial-wave expansion. In the latter case only the s-wave ($\ell = 0$) yields a nonvanishing contribution. Thus, the ejected electron is in the s state with energy

$$E = 2E_{2s} - E_{1s} = \frac{1}{2} I$$

and momentum

$$p = \sqrt{2mE} = \frac{\eta}{\sqrt{2}} .$$

For such an electron,

$$\langle \psi_p | V_{i\mu} | 0 \rangle = 4\pi N_p \varphi(\mu), \quad N_p^2 = 2\pi\xi, \qquad (33)$$

$$\varphi(\mu) = \frac{(\mu + ip)^{i\xi - 1}}{(\mu - ip)^{i\xi + 1}}, \quad \xi = \frac{\eta}{p} = \sqrt{2}, \qquad (34)$$

$$M_A = \sqrt{2}\, 4\pi N_p \hat{B} [\varphi(\nu) - \varphi(\eta_2)]. \qquad (35)$$

After all the derivatives in $\hat{B}$ have been calculated, we are left with terms that depend on $\varphi(\nu)$. Using the chain of equalities

$$\left(\frac{\nu + ip}{\nu - ip} \right)^{i\xi} = \exp\left(-2\xi \arctan \frac{p}{\nu} \right) = \exp\left(-2\sqrt{2} \arctan \frac{1}{2\sqrt{2}} \right) \equiv (\exp), \qquad (36)$$

we can write the final expression for M_A as

$$M_A = 4\pi\sqrt{2}\, N \left(\frac{2}{3} \right)^9 \frac{(\exp)}{\eta^5}, \quad N = N_p N_1 N_2^2, \qquad (37)$$

$$M_A^2 = \frac{\sqrt{2}}{\eta} (\exp)^2 \left(\frac{4}{9} \right)^9 . \qquad (38)$$

The differential and total Auger-decay rates assume the form

$$dw_A=(4\pi\alpha)^2M_A^2\frac{d^3p}{(2\pi)^3}2\pi\delta(E+E_{1s}-2E_{2s})$$

$$=4m\alpha^2(\exp)^2\left(\frac{4}{9}\right)^9 d\Omega_p \tag{39}$$

($d\Omega_p$ is the solid angle into which the final electron lands), and

$$w_A=16\pi m\alpha^2(\exp)^2\left(\frac{4}{9}\right)^9. \tag{40}$$

The expression (39) for dw_A shows that the ionized electrons are distributed isotropically, as they should be for the s-wave. If we wish to find the decay rate in ordinary units (s^{-1}), we need to substitute $m\alpha^2=4.134\times10^{16}\ \mathrm{s}^{-1}$ into (40). This yields

$$w_A=2.06\times10^{14}\ \mathrm{s}^{-1}. \tag{41}$$

The Auger width of the level in electron volts ($m\alpha^2=27.212$ eV) is

$$\Gamma_A=0.135\ \ \mathrm{eV}. \tag{42}$$

The autoionization width of the $2s^2$ level is calculated in Ref. 26, where its value is 0.140 eV. Another very close value is given in Refs. 27 and 5: $\Gamma_A=0.138$ eV. However, such good agreement is observed only for helium. The result (42) is independent of Z, while the widths obtained in the diagonalization approximation[26] increase with Z and reach their asymptotic value of 0.226 eV by $Z=6$. Thus, the discrepancy between (42) and the results of more accurate calculation does not exceed 50% for any value of Z. The same accuracy can probably be guaranteed for the results of the calculations of the excitation cross section of $2s^2$ states of helium-like ions done in the present paper.

The authors would like to express their gratitude to V. G. Gorshkov for fruitful discussions.

[*)] E-mail: mikhailo@thd.pnpi.spb.ru

[1)] The one-electron transition cross section is independent of Z if the photon energy is expressed in units of η.[28]

[1] T. A. Carlson, Phys. Rev. **156**, 142 (1967).
[2] F. W. Byron, Jr. and C. J. Joachain, Phys. Rev. **164**, 1 (1967).
[3] T. Aberg, Phys. Rev. A **2**, 2 (1970); **2**, 1736 (1970).
[4] M. Ya. Amusia, E. G. Drukarev, V. G. Gorshkov, and M. P. Kazachkov, J. Phys. B **8**, 1248 (1975).
[5] J. P. Giese, M. Schilz, J. K. Swenson *et al.*, Phys. Rev. A **42**, 1231 (1990).
[6] J. A. R. Samson, C. H. Greene, and R. J. Bartlett, Phys. Rev. Lett. **71**, 201 (1993).
[7] L. R. Andersson and Burgdörfer, Phys. Rev. Lett. **71**, 50 (1993).
[8] J. H. McGuire, N. Berah, R. J. Bartlett *et al.*, J. Phys. B **28**, 913 (1995).
[9] T. Ishihara, K. Hino, and J. H. McGuire, Phys. Rev. A **44**, R698 (1991).
[10] A. Dalgarno and H. R. Sadeghpour, Phys. Rev. A **46**, R3591 (1992).
[11] J. C. Levin, D. W. Lindle, N. Keller *et al.*, Phys. Rev. Lett. **67**, 968 (1991).
[12] J. C. Levin, I. A. Sellin, B. M. Johnson *et al.*, Phys. Rev. A **47**, R16 (1993).
[13] L. R. Andersson and J. Burgdörfer, Phys. Rev. A **50**, R2810 (1994).
[14] K. Hino, P. M. Bergstrom, Jr., and J. H. Macek, Phys. Rev. Lett. **72**, 1620 (1994).
[15] T. Surić, K. Pisk, B. A. Logan, and R. H. Pratt, Phys. Rev. Lett. **73**, 790 (1994).
[16] M. Ya. Amusia and A. I. Mikhailov, Phys. Lett. A **199**, 209 (1995); J. Phys. B **28**, 1723 (1995).
[17] J. C. Levin, G. B. Armen, and I. A. Sellin, Phys. Rev. Lett. **76**, 1220 (1996).
[18] L. Spielberger, O. Jagutzki, B. Krässig, U. Meyer, Kh. Khayyat, V. Mergel, Th. Tschentscher, Th. Buslaps, H. Brauning, R. Dorner, V. Vogt, M. Achler, J. Ullrich, D. S. Ullrich, D. S. Gemmell, and H. Schmidt-Bocking, Phys. Rev. Lett. **76**, 4685 (1996).
[19] T. Surić, K. Pisk, and R. H. Pratt, Phys. Lett. A **211**, 289 (1996).
[20] R. L. Brown, Phys. Rev. A **1**, 341 (1970).
[21] M. Ya. Amus'ya and A. I. Mikhaĭlov, Zh. Éksp. Teor. Fiz. **111**, 862 (1997) [JETP **84**, 474 (1997)].
[22] J. O. P. Pedersen and P. Hvelplund, Phys. Rev. Lett. **62**, 2373 (1989).
[23] J. C. Straton, J. H. McGuire, and Z. Chen, Phys. Rev. A **46**, 5514 (1992).
[24] V. A. Sidorovich, Zh. Éksp. Teor. Fiz. **103**, 806 (1993) [JETP **76**, 396 (1993)].
[25] V. A. Sidorovich, Physica Scripta **50**, 119 (1994).
[26] U. I. Safronova and V. S. Senashenko, Opt. Spektrosk. **45**, 9 (1978) [Opt. Spectrosc. **45**, 4 (1978)].
[27] W. Shearer-Izumi, At. Data Nucl. Data Tables **20**, 531 (1977).
[28] V. G. Gorshkov, A. I. Mikhaĭlov, and S. G. Sherman, Zh. Éksp. Teor. Fiz. **66**, 2020 (1974) [Sov. Phys. JETP **39**, 995 (1974)].
[29] V. G. Gorshkov and V. S. Polikanov, JETP Lett. **9**, 484 (1969) [*sic.*].
[30] A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics*, Wiley, New York (1974).
[31] V. G. Gorshkov, A. I. Mikhaĭlov, and V. S. Polikanov, Nucl. Phys. **55**, 273 (1964).

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

High-order perturbation theory for the hydrogen atom in a magnetic field

V. M. Vaĭnberg,*) V. A. Gani,†) and A. E. Kudryavtsev‡)

Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia

(Submitted 12 May 1997)
Zh. Éksp. Teor. Fiz. **113**, 550–562 (February 1998)

The states of a hydrogen atom with principal quantum numbers $n \leqslant 3$ in a constant uniform magnetic field $\mathscr{H}$ are studied. Coefficients in the expansion of the energy of these states in powers of $\mathscr{H}^2$ up to the 75th order are obtained. Series for the energies of the states and the wave functions are summed to values of $\mathscr{H}$ on the order of the atomic magnetic field. A generalization of the moment method upon which these calculations are based can be used in other cases in which a hydrogen atom is perturbed by a potential with a polynomial dependence on the coordinates. © *1998 American Institute of Physics.* [S1063-7761(98)01202-5]

1. INTRODUCTION

A new aspect of the problem of the hydrogen atom in constant external electric $\mathscr{E}$ and magnetic $\mathscr{H}$ fields has recently been discovered. It has been noticed[1] that the asymptotic behavior of the perturbation series expansion in $\mathscr{E}$ changes significantly for certain values of $\mathscr{H}$. This change is related to the influence of previously neglected complex solutions of the classical equations of motion on the asymptotic behavior. A look at this phenomenon from the standpoint of the expansion in terms of $\mathscr{H}$ might be useful. Our work is oriented specifically in this direction; here we propose an efficient method for constructing the perturbation series, and we discuss the asymptotic behavior of this series in the Zeeman effect.

A moment method has been proposed[2] for calculating the higher orders of perturbation theory, which has a number of additional possibilities compared to the conventional recurrence methods. It was described as a means for studying the $1/n$-expansion (dimensional expansion) in the problem leading to an effective isotropic anharmonic oscillator. Later, it was applied to a $1/n$-expansion for three interacting bodies[3] in which the effective anharmonic oscillator was anisotropic.

We note that the perturbation of the hydrogen atom by a polynomial potential can also be conveniently studied by the moment method. Uniform electric and magnetic fields are perturbations of this kind. The advantages of the moment method are clearly evident in the problem of the Zeeman effect. High orders of perturbation theory for this problem have been studied elsewhere[4–6]; 36 coefficients in the expansion of the ground state energy of the hydrogen atom in powers of $\mathscr{H}^2$ have been published.[6] For the excited states, no Zeeman coefficients in the perturbation theory have been presented beyond the third order in $\mathscr{H}^2$.[7–10] In the Schrödinger equation for the Zeeman effect the variables do not separate, and this makes it difficult to calculate the higher orders of perturbation theory. The moment method does not require separation of variables. In addition, as will be shown here, this method can be applied to degenerate states. This is obviously an important possibility for most states of the hydrogen atom.

Logarithmic perturbation theory has probably been the most frequently used for recursive calculations of the perturbation theory coefficients.[11–15] Because of its simple algebraic structure, logarithmic perturbation theory has made it possible to calculate corrections to the highest orders in some cases. For example, 160 orders of perturbation theory for the Stark shift of the ground state of the hydrogen atom have been obtained in this way.[16] Logarithmic perturbation theory is not, however, free of limitations. Even a single node in the wave function leads to a substantially more complicated computational scheme.[17] Logarithmic perturbation theory looks much more complicated when the variables are not separable. Without separation of variables, it has only been used to calculate a few initial orders of perturbation theory for the hydrogen atom in electric and magnetic fields.[9,10,14]

Here it is worth recalling an old variant of perturbation theory based on the generalized virial theorem and the Hellman–Feynman theory, which may be regarded as the predecessor of the modern method of moments. Recurrence relations have been written down for spherically symmetric problems[17,18] which make it as easy to calculate the corrections to the energies of states with nodes as to those without nodes. However, the moments given in those papers were diagonal, and this narrowed the range of applicability of the method. For example, it has not been possible to recover the wave function using diagonal moments.

Theoretical approaches to studying the behavior of the hydrogen atom in constant electric and magnetic fields are reviewed by Lisitsa.[20]

In this paper we derive recurrence relations by Ader's moment method for the hydrogen atom in a uniform magnetic field. Two examples illustrate how to apply them in the case of a nondegenerate state and when degeneracy is present. Then we examine the asymptotic behavior of the resulting numerical expansion coefficients for the level energies and present the result of summing these perturbation series. With the ground state as an example, it is shown how the wave function of the perturbed hydrogen atom can be obtained from the moments.

2. BASIC RECURRENCE RELATION

Let us consider the state $|\psi_0\rangle = |n,l\rangle$ of the hydrogen atom with principal quantum number n, angular momentum l, and zero projection of the momentum $m=0$, perturbed by a strong magnetic field $\mathscr{H}=\mathscr{H}_z$. We expand the energy of this state and its wave function in the form

$$E=\sum_{k=0}^{\infty} E_k \gamma^{2k}, \quad \psi(\mathbf{r})=\sum_{k=0}^{\infty} \psi_k(\mathbf{r})\gamma^{2k}, \tag{1}$$

where

$$\gamma = n^3 \mathscr{H}/\mathscr{H}_0, \quad \mathscr{H}_0 = e^3 m^2 c/\hbar^3 = 2.35\cdot 10^9 \text{ G}.$$

Here ψ_k is the correction of order k to the Coulomb wave function that satisfies the inhomogeneous equation

$$(\hat{H}_0 - E_0)\psi_k = -\hat{H}_1 \psi_{k-1} + \sum_{j=1}^{k} E_j \psi_{k-j}, \tag{2}$$

where

$$\hat{H}_0 = -\frac{1}{2}\nabla^2 - \frac{1}{r}, \quad \hat{H}_1 = \frac{1}{8}(r^2 - z^2).$$

We use atomic units.

In order to move from the differential equation to an algebraic equation, we introduce the moments of order k,

$$P_{\sigma\nu}^{k} = \langle \tilde{\psi}_0 | r^{\sigma-\nu} z^{\nu} | \psi_k \rangle, \tag{3}$$

where $|\tilde{\psi}_0\rangle = C\exp(-r/n)$, and σ and ν are integers. In this definition the wave function $|\tilde{\psi}_0\rangle$ only contains the exponential factor of the unperturbed wave function, which carries its scale and has no nodes. (Note that the overall normalization factor for all the moments can be arbitrary.) Thus, as has been done before,[2] we multiply Eq. (2) on the left by $\langle \tilde{\psi}_0 | r^{\sigma-\nu} z^{\nu}$ and use the fact that the hamiltonian can act to the left on explicitly known functions. As a result, we obtain recurrence relations for the moments of order k:

$$\frac{(\sigma-\nu)(\sigma+\nu+1)}{2} P_{\sigma-2,\nu}^{k} + \frac{\nu(\nu-1)}{2} P_{\sigma-2,\nu-2}^{k} - \frac{\sigma+1-n}{n} P_{\sigma-1,\nu}^{k} + P_{\sigma\nu}^{0} E_k = R_{\sigma\nu}^{k-1}, \tag{4}$$

where

$$R_{\sigma\nu}^{k-1} \equiv \frac{1}{8}\left(P_{\sigma+2,\nu}^{k-1} - P_{\sigma+2,\nu+2}^{k-1}\right) - \sum_{j=1}^{k-1} E_j P_{\sigma\nu}^{k-j}.$$

The right-hand side of Eq. (4) only contains moments of preceding orders. The coefficient E_k, which we refer to as the hypersusceptibility of order k, can also be expressed in terms of moments of preceding orders. This relation follows from Eq. (4) and will be written out separately below for each of the cases to be studied.

Similarly, we can examine another perturbation of the same state if this perturbation has the form of a polynomial in r and z. For this it is sufficient just to change the right-hand side of Eq. (4). The expression in parentheses on the right-hand side, which represents the contribution of the magnetic field, is replaced by another function of the moments of the preceding orders produced by the new perturbation.

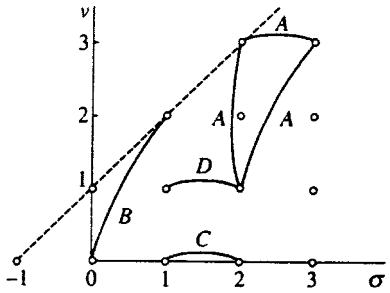


FIG. 1. Subscripts for the moments of order k coupled by the basic recurrence relation. Each of the lines A, B, C, and D represents one of the characteristic cases.

The sequence of calculations relying on Eq. (4) can be made more instructive if we depict points with integer coordinates of the columns σ and rows ν on a lattice plane. The subscripts of the moments required to calculate the corrections to the energies and wave functions lie in the sector $\sigma \geqslant \nu-1$, $\nu \geqslant 0$ of this lattice. In general, Eq. (4) couples moments of order k whose subscripts lie at the vertices of a triangle (see example A in Fig. 1). As a result of the gradual movement of this triangle over the lattice, all the couplings of moments contained in Eq. (4) will appear in sequence. The first term on the left-hand side of Eq. (4) drops out if we set $\sigma=\nu$. Here the triangle (example A) degenerates into a line segment beginning on the dashed line $\sigma=\nu-1$ representing the relation between two moments from different rows (example B). Equation (4) also couples the moments in pairs along each of the lines with $\nu=0$ and $\nu=1$ (examples C and D). In these cases the second term drops out from the left-hand side of Eq. (4). With successive filling from left to right of the rows with $\nu=0$ and $\nu=1$, the calculation of the moments in each order k begins. The initial (left) element of the third row is calculated using Eq. (4) between a pair of moments (see example B in Fig. 1), while subsequent elements of this row are obtained from the ''triangle'' relation (of form A in Fig. 1), etc.

As a direct integration shows, for $k=0$ those moments $\langle \tilde{\psi}_0 | r^{\sigma-\nu} z^{\nu} | n,l \rangle$ for which $l-1 \leqslant \sigma < n-2$ are equal to zero. If the perturbation is even, as in the case of the Zeeman effect, then all the corrections to the wave function have the same parity. Thus, in all orders of perturbation theory those moments for which the sum $\nu+l$ is odd will go to zero.

3. ISOLATED STATES

A magnetic field does not mix states with opposite parities, so besides the ground state, we must treat the $2s$-, $2p$-, and $3p$-states as nondegenerate. We shall show how the moment method works in the latter case. Only moments with odd ν can be nonzero. In the zeroth order,

$$P_{\sigma,2\kappa+1}^{0} = -\frac{\sigma(\sigma+3)!}{18(2\kappa+3)}\left(\frac{3}{2}\right)^{\sigma}, \quad \text{integer } \kappa. \tag{5}$$

In subsequent orders, an additional requirement is imposed: corrections to the wave function $|\psi_0\rangle$ must be orthogonal to it; i.e., $\langle\psi_0|\psi_k\rangle=\delta_{0,k}$. This condition is a standard element of the Brillouin–Wigner perturbation theory,[21] and was used in Ref. 2. In the $3p$-state, it takes the form of the following additional relationship between the momenta:

$$P^k_{11}-\frac{1}{6}P^k_{21}=\delta_{0,k}. \tag{6}$$

In order to obtain a formula for E_k, we first substitute $\nu=1$, $\sigma=1$ into the recurrence relation (4) and then $\nu=1$, $\sigma=2$. The system of two linear equations has the solution

$$E_k=R^{k-1}_{11}-\frac{1}{6}R^{k-1}_{21}=\frac{1}{8}(P^{k-1}_{31}-P^{k-1}_{33})-\frac{1}{48}(P^{k-1}_{41}-P^{k-1}_{43}), \tag{7}$$

with

$$P^k_{01}=5R^{k-1}_{11}-\frac{1}{3}R^{k-1}_{21}. \tag{8}$$

The sum containing the hypersusceptibilities of the previous orders has dropped out of the final expression for E_k because of the orthogonality condition (6).

Equations (4), (6), and (7) form a closed system of recurrence relations. In each order $k\geqslant 1$, the calculations proceed in this sequence: first Eq. (7) is used to calculate E_k. In the next step $\nu=1$ and $\sigma=3$ are substituted into the recurrence relation (4). In this form it, together with the orthogonality condition (6), forms a system of equations that yields the initial elements of the row of moments with $\nu=1$:

$$P^k_{11}=\frac{1}{3}(R^{k-1}_{31}-P^0_{31}E_k)=\frac{1}{6}P^k_{21}.$$

By successively increasing σ by unity, it is easy to arrive at the required limiting moment for this row using Eq. (4). Then substituting $\nu=3$ and $\sigma=3$ into Eq. (4), we obtain the initial moment for the next row:

$$P^k_{33}=3(3P^k_{11}-R^{k-1}_{33}+P^0_{33}E_k)$$

and so on.

The limiting moments, i.e., the moments with the maximum subscripts σ and ν within a given order k, are determined by the following conditions. Calculating the hypersusceptibility of a higher order K requires the moments $P^1_{\sigma\nu}$ in the region $\nu-1\leqslant\sigma\leqslant 3K$, $0\leqslant\nu\leqslant 2K$, the moments $P^2_{\sigma\nu}$ in the region $\nu-1\leqslant\sigma\leqslant 3K-3$, $0\leqslant\nu\leqslant 2K-2$, etc.

It is somewhat simpler to calculate the coefficients E_k for the remaining isolated states.

4. DEGENERATE STATES IN THE MOMENT METHOD

As an example, let us consider the pair of states $|3s\rangle$ and $|3d\rangle$, which are split by a magnetic field. Given the degeneracy, with the aid of a function of the form

$$|\psi_0\rangle=\cos\alpha|3s\rangle+\sin\alpha|3d\rangle, \quad |\widetilde{\psi}_0\rangle\sim e^{-r/3},$$

it is easy to obtain the zeroth-order moments:

$$P^0_{\sigma,2\kappa}=\frac{(\sigma+2)!}{54(2\kappa+1)}\left(\frac{3}{2}\right)^{\sigma}\left[\sigma(\sigma+1)-\frac{\kappa(\sigma+3)(\sigma+4)}{2\kappa+3}\xi\right], \tag{9}$$

where

$$\xi=\sqrt{2}\tan\alpha.$$

In all orders of perturbation theory, the moments in the odd rows are equal to zero; i.e., $P^k_{\sigma,2\kappa+1}=0$. The condition that the corrections be normal to the wave function in its zeroth approximation is equivalent to the following relation between the moments:

$$12P^k_{10}-\frac{1}{3}(4+\xi)P^k_{20}+\xi P^k_{22}=18P^k_{00}, \quad k\geqslant 1. \tag{10}$$

The hypersusceptibility E_k can be expressed in terms of the moments of preceding orders in two independent ways.

a) The moment $P^k_{-1,0}$, needed in the next step for the system of linear equations, can be obtained from Eq. (4) with $\nu=0$ and $\sigma=0$. Note that the coefficient E_k dropped out of Eq. (4) because $P^0_{00}=0$. A system of equations containing E_k results if we first set $\nu=0$, $\sigma=1$ and then $\nu=0$, $\sigma=2$ in Eq. (4). Its solution is

$$E^{(a)}_k=\frac{9}{2}R^{k-1}_{00}-3R^{k-1}_{10}+\frac{1}{3}R^{k-1}_{20}, \tag{11a}$$

and

$$P^k_{00}=-9R^{k-1}_{00}+6R^{k-1}_{10}-\frac{1}{3}R^{k-1}_{20}. \tag{12}$$

The moment P^k_{00} is obtained along the way, and is substituted into the right-hand side of the orthogonality condition (10).

b) Substituting $\nu=2$ and $\sigma=2$ into Eq. (4), as well as the value found for P^k_{00}, we obtain a second independent expression for E_k:

$$E^{(b)}_k=\frac{1}{2}\left(9R^{k-1}_{00}-6R^{k-1}_{10}+\frac{1}{3}R^{k-1}_{20}+R^{k-1}_{22}\right). \tag{11b}$$

Equations (11a) and (11b) yield a common value of the magnetic susceptibility E_1 for two values of ξ:

$$\xi=\xi_{1,2}=\frac{-13\pm 3\sqrt{41}}{10}. \tag{13}$$

This is a natural result that follows from the secular equation in the Rayleigh–Schrödinger perturbation theory.

In the following the notation $3s$ is retained for the state with a small admixture of d-wave and $\xi=\xi_1=(3\sqrt{41}-13)/10$, while $3d$ denotes the state orthogonal to it, a mixture of s- and d-waves. In the subsequent approximations, the uniqueness condition imposed on the hypersusceptibility of the $(k+1)$-st order, $E^{(a)}_{k+1}=E^{(b)}_{k+1}$, is equivalent to an equation coupling seven unknown moments of order k. There is still another relation between the moments, the orthogonality condition (10). In order to obtain a closed system, the orthogonality and uniqueness conditions must be

TABLE I. Hypersusceptibilities of the hydrogen atom in a magnetic field.

k	E_k		
	$1s$-state	$2s$-state	$2p$-state
1	$+2.5000\cdot 10^{-1}$	$+3.5000\cdot 10^{0}$	$+1.5000\cdot 10^{0}$
2	$-2.7604\cdot 10^{-1}$	$-1.5933\cdot 10^{2}$	$-4.2000\cdot 10^{1}$
3	$+1.2112\cdot 10^{0}$	$+2.2508\cdot 10^{4}$	$+4.2400\cdot 10^{3}$
4	$-9.7554\cdot 10^{0}$	$-5.5166\cdot 10^{6}$	$-7.4365\cdot 10^{5}$
5	$+1.8630\cdot 10^{2}$	$+1.8817\cdot 10^{9}$	$+1.8710\cdot 10^{8}$
6	$-1.9593\cdot 10^{3}$	$-8.2044\cdot 10^{11}$	$-6.2321\cdot 10^{10}$
7	$+4.2749\cdot 10^{4}$	$+4.3896\cdot 10^{14}$	$+2.6362\cdot 10^{13}$
8	$-1.1869\cdot 10^{6}$	$-2.8169\cdot 10^{17}$	$-1.3804\cdot 10^{16}$
9	$+4.0973\cdot 10^{7}$	$+2.1373\cdot 10^{20}$	$+8.7878\cdot 10^{18}$
10	$-1.7252\cdot 10^{9}$	$-1.8979\cdot 10^{23}$	$-6.7029\cdot 10^{21}$
25	$+4.9366\cdot 10^{40}$	$+5.3146\cdot 10^{73}$	$+6.0276\cdot 10^{71}$
50	$-3.1317\cdot 10^{109}$	$-1.8694\cdot 10^{173}$	$-9.9461\cdot 10^{170}$
75	$+3.3150\cdot 10^{189}$	$+5.8114\cdot 10^{283}$	$+2.0188\cdot 10^{281}$
	$3s$-state	$3p$-state	$3d$-state
1	$+1.9579\cdot 10^{1}$	$+9.0000\cdot 10^{0}$	$+5.1715\cdot 10^{0}$
2	$-7.9926\cdot 10^{3}$	$-3.5311\cdot 10^{3}$	$-1.0174\cdot 10^{3}$
3	$+9.8654\cdot 10^{6}$	$+3.4481\cdot 10^{6}$	$+6.4414\cdot 10^{5}$
4	$-2.0927\cdot 10^{10}$	$-5.4496\cdot 10^{9}$	$-8.6478\cdot 10^{8}$
5	$+5.8826\cdot 10^{13}$	$+1.1757\cdot 10^{13}$	$+1.6415\cdot 10^{12}$
6	$-2.0349\cdot 10^{17}$	$-3.1806\cdot 10^{16}$	$-4.1602\cdot 10^{15}$
7	$+8.3124\cdot 10^{20}$	$+1.0348\cdot 10^{20}$	$+1.3551\cdot 10^{19}$
8	$-3.9194\cdot 10^{24}$	$-3.9589\cdot 10^{23}$	$-5.4574\cdot 10^{22}$
9	$+2.1055\cdot 10^{28}$	$+1.7580\cdot 10^{27}$	$+2.6300\cdot 10^{26}$
10	$-1.2786\cdot 10^{32}$	$-8.9851\cdot 10^{30}$	$-1.4785\cdot 10^{30}$
25	$+1.3794\cdot 10^{94}$	$+2.5255\cdot 10^{92}$	$+1.0620\cdot 10^{92}$
50	$-9.3229\cdot 10^{211}$	$-7.3330\cdot 10^{209}$	$-6.1021\cdot 10^{209}$
75	$+2.8054\cdot 10^{340}$	$+1.3961\cdot 10^{338}$	$+1.7341\cdot 10^{338}$

supplemented by seven more equations that follow from the recurrence relation (4). From the resulting system of nine equations it is sufficient to determine just two moments: P_{10}^{k} and P_{22}^{k}. Then, with the aid of the already known moments and Eq. (4), it is easy to calculate successively all the required moments of the given order, passing row-by-row through the array of subscripts, as in the case of the $3p$-state.

In similar fashion it is possible to do the calculation for a state with arbitrary n. We merely note the key elements of these calculations. The unperturbed wave function has a certain parity and contains g energy-degenerate terms. In each order $k \geqslant 1$ there are two groups of moments. The recurrence relation (4) couples the moments $P_{\sigma\nu}^{k}$ with $\sigma < n-2$ among themselves and, separately, the moments with $\sigma \geqslant n-2$. The moments from the different groups are coupled by the recurrence relations only via moments of preceding orders. In the intermediate stage, g independent expressions for the coefficient E_{k+1} are obtained and the conditions for its uniqueness make it possible to express the moments $P_{\sigma\nu}^{k}$ from the region $\sigma \geqslant n-2$ in terms of the moments from preceding orders.

5. RESULTS

5.1. Energy levels

We have obtained the Zeeman hypersusceptibilities E_k up to the 75-th order for all levels with $n \leqslant 3$ by the moment method (see Table I). The calculations were carried out to 32 (decimal) significant figures. Complete agreement was found with the results of Ref. 8 and of Ref. 7, which contains the first five coefficients E_k for the ground state and three each of the initial coefficients for the $2s$- and $2p$-states in the form of rational fractions. In Ref. 9 a discrepancy was noted between the equation for the coefficient E_2 obtained there and its value for $l=1$ in Ref. 7. This discrepancy is confirmed here. The formula of Ref. 9 implies $E_2^{(2p)} = -45.556$; our result, $E_2^{(2p)} = -42$, agrees with Ref. 7.

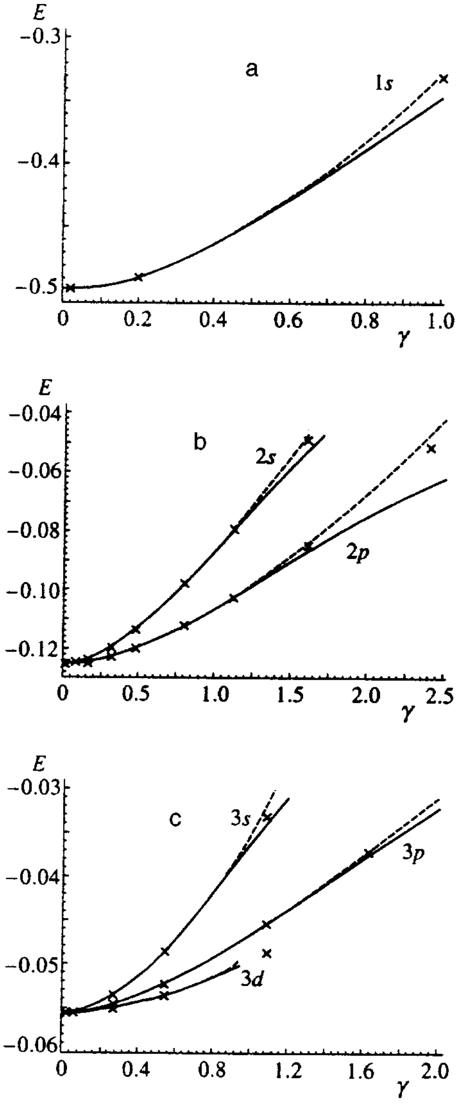


FIG. 2. The sum of perturbation series for the level energies taken using the Padé approximant. Smooth curves $E(\gamma) = [25/25](\gamma^2)$; dashed curves $E(\gamma) = [25/24](\gamma^2)$. The results of Ref. 22 are shown for comparison.

The energy levels $E(\gamma)$ of six states, obtained by summing the corresponding perturbation theory series with the aid of the Padé approximant $[L/L](\gamma^2)$ and $[L/L-1](\gamma^2)$, are shown in Fig. 2. This figure also gives an idea of the convergence region for the Padé approximant. Without expanding in $\mathscr{H}^2$, with the aid of a spline method (one of the modifications of the variational method), the energy eigenvalues of these states are calculated elsewhere[22] for several values of $\mathscr{H}$, and these values are also shown in Fig. 2. Of all the nonperturbative calculations, Ref. 22 has the best accuracy, with eight-place accuracy for the energy levels of the excited states, and it contains a comparison with a large number of previous calculations. The accuracy of the sum of the perturbation series taken with the Padé approximant is fairly high. When $\gamma\approx 1$, 3–4 significant figures are attained for the level energy for the $2s$- and $3p$-states and at least two places for the remaining states. The convergence improves rapidly with decreasing γ, so that for $\gamma\leqslant 0.3$ the accuracy of the perturbation theory sums exceeds that of the variational calculations.[22]

A modification of dimensional scaling has been developed[23] for calculating the Zeeman effect with $\lambda=[1+|m|]^{-1}$ serving as a small parameter. This semiclassical method is convenient in that it is applicable over a wide range of external fields. The energy of the ground state was obtained[23] to an accuracy of $\sim 10^{-5}$, both for $\gamma=0.1$ and $\gamma=1$, by expanding the expansion in λ. Our calculations yield the ground-state energy with an accuracy of $\sim 10^{-13}$ for $\gamma=0.1$ and $\sim 10^{-1}$ for $\gamma=1$.

With increasing order, the perturbation theory coefficients E_k approach an asymptote whose principal term for the Zeeman effect has the form[4,5]

$$\tilde{E}_k=(-1)^{k+1}\frac{D_{nl}}{\pi^{2n+1/2}}\left(\frac{n^2}{\pi}\right)^{2k}\Gamma\left(2k+2n-1+\frac{(-1)^l}{2}\right). \tag{14}$$

For the levels discussed here,

$$D_{1s}=32,\quad D_{2s}=128,\quad D_{2p}=64,$$

$$D_{3s}=\frac{2^{15}}{3^4}\left(\alpha_1-\frac{\alpha_2}{2\sqrt{2}}\right)^2,\quad D_{3d}=\frac{2^{15}}{3^4}\left(\alpha_2+\frac{\alpha_1}{2\sqrt{2}}\right)^2,$$

where

$$\alpha_1=-\left(\frac{1}{2}+\frac{13}{6\sqrt{41}}\right)^{1/2},\quad \alpha_2=\left(\frac{1}{2}-\frac{13}{6\sqrt{41}}\right)^{1/2}.$$

This result was obtained using the methods employed by Bender and Wu.[24] First the barrier penetration was calculated in the semiclassical approximation for an imaginary value of the magnetic field, then the dispersion relation in terms of $\mathcal{H}^2$ was applied. According to the conditions of Ref. 5, strictly speaking Eq. (14) does not apply to the $3p$-state. However, it can be assumed that Eq. (14) encompasses all six states being discussed here, and the coefficient D_{3p} can be selected by comparing E_k and $\tilde{E}_k$ for the higher orders. This yields $D_{3p}=2^{13}/3^3$. The approach of the exact coefficients E_k to the asymptote (14) is illustrated in Fig. 3.

Corrections to the asymptotic $\tilde{E}_k$ for several states, but not $2s$, $2p$, and $3p$, were obtained in Refs. 4 and 5. Writing the corrections in the form

$$\frac{E_k}{\tilde{E}_k}=c_0+\frac{c_1}{2k}+\frac{c_2}{(2k)^2}+\dots, \tag{15}$$

it is easy, following the method of Ref. 24, to find the coefficients c_i (see Table II). In all cases considered here, $|c_0-1|<10^{-6}$. As the number of corrections calculated in Eq. (15) increases, the accuracy of the coefficients c_i increases. The stability of the values of the power series corrections on going to a large number of these corrections confirms the validity of the determination of the principal term of the $\tilde{E}_k$ for the $3p$-state. Recall that by virtue of the dispersion relation in $\mathcal{H}^2$,[5] the coefficients c_i are related to the corrections to the semiclassical approximation for the barrier penetration when $\mathcal{H}^2<0$. Direct calculation of the semiclassical corrections is a tedious task.

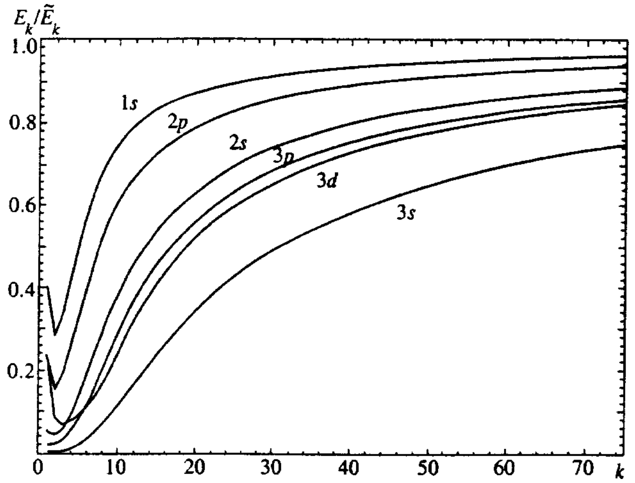


FIG. 3. The approach of the exact hypersusceptibilities E_k to the corresponding asymptotes $\tilde{E}_k$ for six states of the hydrogen atom.

5.2. Wave functions

By analogy with the anharmonic oscillator,[2] the correction $|\psi_k\rangle$ to the Coulomb wave function has the form of a polynomial (in r and $\cos\theta$) multiplied by the exponential $|\tilde{\psi}_0\rangle$. The perturbation is polynomial, but the operator on the left-hand side of Eq. (2) does not alter the indicated structure of the wave function $|\psi_k\rangle$. For the ground state

$$|\psi_k\rangle=\left(\sum_{j=0}^{k}\sum_{i=2j}^{3k}a_{ij}^{(k)}r^i\cos^{2j}\theta\right)|\tilde{\psi}_0\rangle \equiv\left(\sum_{j=0}^{k}\sum_{i=2j}^{3k}a_{ij}^{(k)}r^{i-2j}z^{2j}\right)|\tilde{\psi}_0\rangle. \tag{16}$$

Let us clarify the origin of the limits on the sums. The dependence on the angle θ comes only from the expression $r^2\cos^2\theta$ in $\hat{H}_1$, so that the minimum power of the radius in the inner sum of Eq. (16) coincides with the power of $\cos\theta$, while the maximum power of $\cos^2\theta$ coincides with the order of perturbation theory. Using Eq. (2), we can verify that the

TABLE II. Coefficients c_i in the power-series corrections to the asymptotic $\tilde{E}_k$ for the hypersusceptibilities of the hydrogen atom.

State	c_1	c_2	c_3	c_4
$1s$	-2.61829	$+1.282$	-2.6	-11
$2s$	-8.938	$+37.44$	-121	$+2.7\cdot 10^2$
$2p$	-4.6065	$+8.24$	-14.3	-4
$3p$	-11.227	$+59.5$	-239	$+6\cdot 10^2$

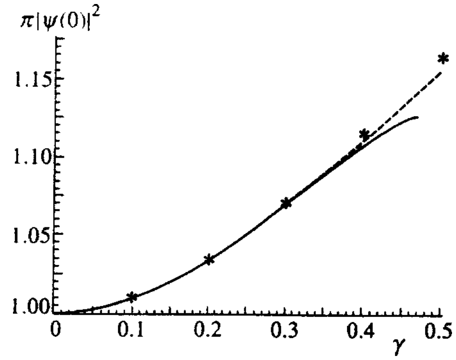


FIG. 4. Ground-state wave function calculated by the moment method. The smooth curve was obtained using the Padé approximant $[9/9](\gamma^2)$, the dashed, using $[9/8](\gamma^2)$. Asterisks indicate the results of Ref. 25.

maximum power of the radius in $|\psi_k\rangle$ is three units greater than in $|\psi_{k-1}\rangle$. Equation (16) yields a system of linear equations for the coefficients $a^{(k)}_{ij}$:

$$\sum_{j=0}^{k}\sum_{i=2j}^{3k} P^0_{i+\alpha,j+\beta}a^{(k)}_{ij}=P^k_{\alpha\beta},$$

$$2\beta\leqslant\alpha\leqslant 3k,\quad 0\leqslant\beta\leqslant k. \qquad (17)$$

Instead of the values of the subscripts α and β chosen here, it is possible to choose another set, which yields $(2k+1)(k+1)$ independent values. This possibility is useful for checking the accuracy of the calculations. We also verified the orthogonality of the corrections (16) to the wave functions $|\psi_0\rangle$. In our calculations the orthogonality was retained with reasonable accuracy up to the 18th order of perturbation theory. The corrections to the wave functions up to the second order are given in Ref. 7. On comparing them with our results, we noticed only one discrepancy. The coefficient of $r^3\gamma^2$ in Ref. 7 has an incorrect sign and this causes a nonorthogonality in the first order correction to the unperturbed wave function in that paper.

We calculated the value of $|\psi(0)|^2$ by the moment method, using the Padé approximant to sum the perturbation series for the normalization factor and for the wave function itself. The results are shown in Fig. 4. It is clear that with 18 orders of perturbation theory it is possible to advance to $\gamma\approx 0.4$. Here agreement is maintained with an earlier calculation[25] of the ground-state wave function by other methods with an accuracy of $\sim 10^{-3}$ for the magnetic fields of interest to us.

6. CONCLUSION

The good agreement between energy eigenvalues obtained by summing perturbation series and the corresponding results of independent variational calculations serve, on one hand, to confirm the validity of calculating the perturbation coefficients themselves, and on the other, to show that there are no nonperturbative contributions to the energy of these states (the summation technique has been chosen correctly).

The study described here demonstrates the high efficiency of the Ader moment method. Here we note the features of this method that have made it possible to obtain high orders of perturbation theory for the Zeeman effect. These features can be useful in solving other perturbation theory problems.

1) Simple recurrence relations are obtained without separation of variables in the original equation. With their help the procedure for calculating the corrections becomes purely algebraic.

2) Nodes of the wave function are not explicitly incorporated in the moment method, and do not make the calculations more complicated. This makes it possible to consider excited states.

3) Degeneracy of the levels, as demonstrated here for the first time, does not preclude the use of the moment method, although it does lead to a more cumbersome procedure for solving the recurrence relations.

4) In addition, the set of moments obtained while determining the hypersusceptibilities contains information on the wave function. Corrections to the Coulomb wave function, like the corrections to the energy levels, can be obtained purely algebraically.

We have limited ourselves to the case $m=0$ here. Obviously, for any state with nonzero projection of the angular momentum, the paramagnetic part of the perturbation, $0.5\gamma\hat{L}_z$, owing to its diagonal nature, simply shifts the energy by $(m/2)\gamma$ and does not change the wave function. The effect of the diamagnetic part of the perturbation, $\gamma^2\hat{H}_1$, on this state can be calculated by the moment method, relying on a recurrence relation analogous to Eq. (4).

As already noted, application of the moment method to the hydrogen atom is not restricted to the Zeeman effect. Any perturbation of the form

$$V(r)=\sum b_{ij}r^{i-j}z^j$$

yields the recurrence relation (4) with an obvious simple change in the right hand side. Thus, for example, it is possible to include the combined effect of external electric and magnetic fields that are uniform or take the form of a finite sum of multipoles.

In conclusion, we would like to express our deep gratitude to V. S. Popov, who read the manuscript and made a number of valuable comments which led to improvements. We also thank V. G. Ksenzov for useful discussions and V. D. Mura and B. M. Karnakov for their interest in this work. One of the authors (V.A.G.) is indebted to V. L. Morgunov and A. A. Panfilov for discussing several questions related to the numerical calculations.

[*] E-mail: wein@vitep5.itep.ru

[†] Moscow Institute of Engineering Physics, e-mail: gani@vitep5.itep.ru

[‡] E-mail: kudryavtsev@vitep5.itep.ru

---

[1] V. S. Popov and A. V. Sergeev, JETP Lett. **63**, 417 (1996).
[2] J. P. Ader, Phys. Lett. A **97**, 178 (1983).
[3] D. Z. Goodson and D. R. Herschbach, Phys. Rev. Lett. **58**, 1628 (1987).
[4] B. G. Adams, J. E. Avron, J. Čižek *et al.*, Phys. Rev. A **21**, 1914 (1980).
[5] J. E. Avron, Ann. Phys. (N.Y.) **131**, 73 (1981).
[6] J. Čižek and E. R. Vrscay, Int. J. Quantum Chem. **21**, 27 (1982).
[7] A. Galindo and P. Pascual, Nuovo Cimento B **34**, 155 (1976).

[8] B. R. Johnson, K. F. Scheibner, and D. Farrelly, Phys. Rev. Lett. **51**, 2280 (1983).
[9] A. V. Turbiner, Z. Phys. A **308**, 111 (1982).
[10] A. V. Turbiner, Zh. Éksp. Teor. Fiz. **84**, 1329 (1983) [Sov. Phys. JETP **57**, 770 (1983)].
[11] V. S. Polikanov, Zh. Éksp. Teor. Fiz. **52**, 1326 (1967) [Sov. Phys. JETP **25**, 882 (1967)].
[12] V. S. Pekar, Teor. Mat. Fiz. **9**, 140 (1971).
[13] A. D. Dolgov and V. S. Popov, Phys. Lett. B **86**, 185 (1979).
[14] Y. Aharonov and C. K. Au, Phys. Rev. A **20**, 2245 (1979); **22**, 328 (1980).
[15] V. Privman, Phys. Rev. A **22**, 1833 (1980).
[16] V. L. Eletskiĭ and V. S. Popov, Dokl. Akad. Nauk SSSR **250**, 74 (1980); S. P. Alliluev, V. L. Eletsky, and V. S. Popov, Phys. Lett. A **73**, 103 (1979).
[17] S. P. Alliluev, V. M. Vaĭnberg, V. L. Eletskiĭ, and V. S. Popov, Zh. Éksp. Teor. Fiz. **82**, 77 (1982) [Sov. Phys. JETP **55**, 46 (1982)].
[18] R. J. Svenson and S. H. Danforth, J. Chem. Phys. **57**, 1734 (1972).
[19] J. Killingbeck, Phys. Lett. A **65**, 87 (1978).
[20] V. S. Lisitsa, Usp. Fiz. Nauk **153**, 379 (1987) [Sov. Phys. Usp. **30**, 927 (1987)].
[21] J. Ziman, *Modern Quantum Theory* [Russian translation], Mir, Moscow (1971), Ch. 3.
[22] Jang-Huar Wang and Chen-Shiung Hsue, Phys. Rev. A **52**, 4508 (1995).
[23] C. M. Bender, L. D. Mlodinov, and N. Papanicolaou, Phys. Rev. A **25**, 1305 (1982).
[24] C. M. Bender and T. T. Wu, Phys. Rev. D **7**, 1620 (1973).
[25] D. Cabib, E. Fabri, and G. Fiorio, Nuovo Cimento **10**, 185 (1972).

Note added in proof (1 December 1997); We have released a more detailed table of hypersusceptibilities E_k as an electronic preprint, Physics/9708005.

Translated by D. H. McNeill

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Scattering of atoms in the field of counterpropagating light waves. Effect of initial conditions

V. I. Romanenko and L. P. Yatsenko

Institute of Physics, Ukrainian National Academy of Sciences, 252650 Kiev, Ukraine

(Submitted 23 May 1997)
Zh. Éksp. Teor. Fiz. **113**, 563–572 (February 1998)

The scattering of an atom in the field of counterpropagating light waves is studied under conditions such that the state of the atom is a superposition of the ground and excited states. For the case in which this superposition is created by the field of a traveling wave, the momentum distribution function of the atom after scattering by a standing wave is found analytically in the approximation of a short interaction time, when the atom's motion can be neglected. Longer interactions of the atom with the field are studied numerically. We also consider the case of counterpropagating light waves consisting of Gaussian or supergaussian pulses. © *1998 American Institute of Physics.* [S1063-7761(98)01302-X]

1. The motion of atoms in the field of standing light waves has been the subject of thorough study for many years.[1–6] This is because of the interesting physics of the multiphoton processes that occur, as well as of the widespread applications of standing waves in various atomic optics devices.[7] It should be noted that essentially all papers devoted to this topic have been concerned with the scattering of atoms that are initially in the ground state. We have shown[8] that this yields a substantially depleted picture of the possible motion. It turns out that if the atom is initially in a coherent superposition of the ground and excited states that has been prepared in a certain way, then a new effect occurs—an asymmetry in the scattering of the atom by a standing wave field. This phenomenon may serve as a basis for explaining the experimentally observed[9,10] scattering pattern of atoms in the field of short, counterpropagating laser pulses.

In this paper we analyze in detail the scattering of an atom in the field of a standing light wave in the case where, at the beginning of the interaction with the standing light wave, the atom is in a coherent superposition of the ground and excited states. In particular, we examine the interaction of the atom with a standing light wave in the parameter range where the Doppler effect is significant. Numerical simulations are also made of the interaction of atoms with the field of two counterpropagating waves whose envelopes have a time dependence close to the time dependences of the light waves interacting with an atom in an experiment (Gaussian and ''supergaussian'' pulses that are shifted in time). We assume that the time for the interaction of the atom with the light fields is short, so that spontaneous emission from the upper level can be neglected over the time in which the fields act. In addition, as our goal is to clarify the qualitative effect of the initial spatial modulation in the probability amplitudes of the ground or excited states on the scattering pattern, in this paper we have restricted ourselves to a two-level model for the atom.

2. The Schrödinger equation describing the interaction of a two-level atom with an electromagnetic wave field has the form

$$i\hbar\dot{\Psi}=(H_0-\hat{\mathbf{d}}\cdot\mathbf{E})\Psi, \tag{1}$$

where

$$H_0=\frac{\hat{\mathbf{p}}^2}{2m}+\hbar\omega_0|e\rangle\langle e|, \tag{2}$$

$\hat{\mathbf{p}}=-i\hbar\nabla$ is the momentum operator for the atom, m is its mass, $\hbar\omega_0$ is the eigenvalue of the Hamiltonian of the motionless atom in the absence of the light fields that corresponds to the excited state $|e\rangle$ of the atom. (We assume that the eigenvalue corresponding to the ground state, $|g\rangle$, equals zero.), $\hat{\mathbf{d}}$ is the dipole moment operator, and $\mathbf{E}$ is the electric field strength of the light wave.

The interaction of an atomic beam with a standing wave field is known to cause it to spread out in momentum space.[11,12] The magnitude of this spreading increases linearly with time until the offset of the atomic transition from the field frequency caused by the Doppler effect reduces the efficiency of the interaction of the atoms with the field.

We first examine the scattering of an atom in a standing wave field for short interaction times between the atom and field, when the Doppler effect can be neglected.

Let an atom interact for $t>0$ with the standing wave field

$$\mathbf{E}=2\mathbf{e}E_0\sin(kx)\cos(\omega t), \tag{3}$$

where $\mathbf{e}$ is the unit vector in the direction of the electric field.

We seek a solution of Eq. (1) in the form

$$\Psi=c_e(x,t)\exp(-i\omega_0 t)|e\rangle+c_g(x,t)|g\rangle. \tag{4}$$

Substituting Eqs. (4) and (3) in Eq. (1), we find equations for $c_g(x,t)$ and $c_e(x,t)$ in the rotating wave approximation:

$$\begin{aligned}\dot{c}_g(x,t)&=iV_0\sin(kx)\exp(i\delta t)c_e(x,t),\\ \dot{c}_e(x,t)&=iV_0\sin(kx)\exp(-i\delta t)c_g(x,t),\end{aligned} \tag{5}$$

where we have introduced the notation $V_0 = \langle e|\hat{\mathbf{d}}\cdot\mathbf{e}|g\rangle E_0/\hbar$ and $\delta=\omega-\omega_0$. We assume that V_0 is real and does not change during the time t_{int} within which the atom interacts with the standing wave field. In deriving Eq. (5), we have neglected the motion of the atom during the time of the interaction with the field and have left out the term $\hat{\mathbf{p}}^2/2m$ in H_0. It is known[7] that the motion of the atom can be neglected for $t_{\text{int}} \ll T_{RN}$ (the Raman–Nath region), where

$$T_{RN} = 2\pi/\sqrt{2\delta_\tau V_0}. \tag{6}$$

Here $\delta_\tau = \hbar k^2/2m$.

Upon solving Eq. (5), we find

$$\begin{aligned} c_g(x,t) = \Big\{ & c_g(x,0)\cos\left(\frac{Vt}{2}\right) + \frac{i}{V}\,[-\delta c_g(x,0) \\ & + 2V_0 \sin(kx) c_e(x,0)]\sin\left(\frac{Vt}{2}\right)\Big\} \exp\left(\frac{i}{2}\,\delta t\right), \\ c_e(x,t) = \Big\{ & c_e(x,0)\cos\left(\frac{Vt}{2}\right) + \frac{i}{V}\,[\delta c_e(x,0) \\ & + 2V_0 \sin(kx) c_g(x,0)]\sin\left(\frac{Vt}{2}\right)\Big\} \\ & \times \exp\left(-\frac{i}{2}\,\delta t\right), \end{aligned} \tag{7}$$

where the notation $V=\sqrt{\delta^2 + 4V_0^2 \sin^2(kx)}$ has been introduced.

We find the average momentum transferred to the atom,

$$\begin{aligned} \langle \Delta p\rangle = -i\hbar \int_{-\infty}^{\infty} \Big[& \Psi^*(x,t)\,\frac{\partial}{\partial x}\,\Psi(x,t) \\ & - \Psi^*(x,0)\,\frac{\partial}{\partial x}\,\Psi(x,0)\Big] dx \end{aligned} \tag{8}$$

up to time t after the onset of the interaction between the atom and the standing wave field. Substituting Eq. (7) in Eq. (8) yields

$$\begin{aligned} \langle \Delta p\rangle = \hbar k V_0 \int_{-\infty}^{\infty} \frac{\cos(kx)}{V^3} \Big\{ & [V^3 t - \delta^2 (Vt - \sin(Vt))] \\ & \times [\rho_{ge}(x) + \rho_{eg}(x)] + 2i\delta V \sin^2\left(\frac{Vt}{2}\right)[\rho_{ge}(x) \\ & - \rho_{eg}(x)] + 2\delta V_0 \sin(kx)[Vt - \sin(Vt)][\rho_{ee}(x) \\ & - \rho_{gg}(x)]\Big\} dx, \end{aligned} \tag{9}$$

where we have introduced the notation

$$\begin{aligned} &\rho_{gg}(x) = c_g^*(x,0) c_g(x,0), \quad \rho_{ee}(x) = c_e^*(x,0) c_e(x,0), \\ &\rho_{ge}(x) = c_g^*(x,0) c_e(x,0), \quad \rho_{eg}(x) = \rho_{ge}^*(x). \end{aligned} \tag{10}$$

At time $t=0$, let the atom be described by a wave packet and let x_0 be some characteristic point in the packet that describes its position in space, such as its ''center of gravity:''

$$x_0 = \int_{-\infty}^{\infty} x[\rho_{gg}(x) + \rho_{ee}(x)]dx. \tag{11}$$

It can be shown that when the dependences of $\rho_{gg}(x)$, $\rho_{ee}(x)$ and $\rho_{ge}(x)$ on the coordinate x can be represented by a function of $x - x_0$ (i.e., the shape of the wave packet of the atom in a system of coordinates attached to its ''center of gravity'' does not depend on its position relative to the nodes and antinodes of the standing wave), the average of Eq. (9) over x_0 yields $\overline{\langle \Delta p\rangle} = 0$. As an illustration, we consider the case of exact tuning of the radiation frequency to the atomic transition frequency ($\delta=0$). Then, for $\langle \Delta p\rangle$ we have

$$\langle \Delta p\rangle = \hbar k V_0 \int_{-\infty}^{\infty} \cos(kx)[\rho_{ge}(x) + \rho_{eg}(x)]\; dx. \tag{12}$$

Upon introducing the notation $\rho_{ge}(x) + \rho_{eg}(x) = f(x - x_0)$ and making the variable substitution $x = x_0 + x'$ in the integral of Eq. (12), we find

$$\langle \Delta p\rangle = \hbar k V_0 \int_{-\infty}^{\infty} \cos(kx' + kx_0) f(x')dx'. \tag{13}$$

Averaging Eq. (13) with respect to x_0 over the interval $(0, 2\pi/k)$ obviously yields $\overline{\langle \Delta p\rangle} = 0$. Thus, when the atomic beam consists of an ensemble of atoms described by identical wave packets initially distributed in a region of width substantially greater than $2\pi/k$, the average momentum transfer to the atomic beam is zero. This conclusion is based on the form of the dependence of the sum $\rho_{ge}(x) + \rho_{eg}(x)$ on x, and x_0. If, for example, we set $\rho_{ge}(x) + \rho_{eg}(x) = f(x - x_0)\cos(kx)$, it is easy to see that $\overline{\langle \Delta p\rangle} \neq 0$. It might be expected that a standing wave could transfer momentum to an atomic beam which is perpendicular to it if $\rho_{ge}(x) + \rho_{eg}(x) = f(x - x_0) f_0(x)$, where $f_0(x)$ is an arbitrary oscillatory function of x with period $2\pi/k$.

In the general case of an arbitrary frequency offset of the light from the atomic transition frequency, $\delta \neq 0$, it can easily be seen from Eq. (9) that the average momentum transferred to an atom may be nonzero if $\rho_{ge}(x)$ and $\rho_{eg}(x)$ contain spatial harmonics with period $2\pi/k$, or if $\rho_{gg}(x)$ or $\rho_{ee}(x)$ contain spatial harmonics with period π/k.

We have proposed[8] creating oscillatory dependences of $\rho_{ge}(x)$ and $\rho_{eg}(x)$ on x with period $2\pi/k$ using a traveling wave that acts for some time on an atom before it interacts with the standing wave field. Oscillatory $\rho_{gg}(x)$ and $\rho_{ee}(x)$ with period π/k can be created when the atoms interact with an auxiliary standing wave field whose nodes are shifted relative to the main standing wave, which ensures a guided change in the momentum of the atoms.[4] Here the first wave causes a spatial modulation in the populations of the ground and excited states, while for $\delta \neq 0$ the second wave causes a guided change in the momentum of the beam atoms.

Here we shall examine the first possibility, keeping the simplicity of realizing it in mind (for example, an atom in the field of counterpropagating pulses or an atomic beam intersecting two parallel spatially displaced and partially overlapping laser beams). In addition, in this case, the atoms do not have to interact for long with the field required to displace them along the standing wave by a distance on the order of the wavelength.

We find $c_g(x,0)$ and $c_e(x,0)$, assuming that for $t<-\tau$ the atom is in the ground state and then interacts for a time τ with the standing wave field

$$\mathbf{E}=\mathbf{e}E_0\sin(kx-\omega t+\varphi). \tag{14}$$

For $\varphi=0$ this field coincides with the field of one of the two travelling wave forming the standing wave (3). As before, we seek a solution of the Schrödinger equation (1) in the form (4). Solving the equations for $c_g(x,t)$ and $c_e(x,t)$,

$$\dot{c}_g(x,t)=-\frac{V_0}{2}c_e(x,t)\exp(-ikx+i\delta t-i\varphi),$$

$$\dot{c}_e(x,t)=\frac{V_0}{2}c_g(x,t)\exp(ikx-i\delta t+i\varphi) \tag{15}$$

with initial conditions $c_g(x,-\tau)=c(x)$ and $c_e(x,-\tau)=0$, we find

$$c_g(x,0)=\left[\cos\frac{\vartheta}{2}-\frac{i\delta}{\tilde{V}}\sin\frac{\vartheta}{2}\right]\exp\left(i\frac{\delta\tau}{2}\right)c(x),$$

$$c_e(x,0)=\frac{V_0}{\tilde{V}}\sin\frac{\vartheta}{2}\exp\left(i\frac{\delta\tau}{2}+ikx+i\varphi\right)c(x), \tag{16}$$

where we have introduce the Rabi frequency $\tilde{V}=\sqrt{\delta^2+V_0^2}$ and the pulse area $\vartheta=\tilde{V}\tau$ of the traveling wave. Equations (16), together with Eqs. (4) and (7), determine the wave function of the atom at times t beginning with its interaction with the standing wave field.

The diffraction pattern of the atomic scattering in the far zone is described by the wave function of the atoms in wave vector K space:

$$f_g(K,t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-iKx}c_g(x,t)dx,$$

$$f_e(K,t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-iKx}c_e(x,t)dx. \tag{17}$$

Here we use the variable K which is related to the momentum p by the formula $K=p/\hbar$. We shall assume that at the beginning of its interaction with the travelling wave field, the atom is described by the probability amplitude

$$c(x)=\frac{1}{\Delta^{1/2}\pi^{1/4}}\exp\left\{-\frac{(x-x_0)^2}{2\Delta^2}\right\}. \tag{18}$$

In the case of exact tuning of the radiation frequency to the atomic transition ($\delta=0$), the integrals in Eq. (17) can be calculated and we find

$$\begin{aligned}f_g(K,t)=\pi^{-1/4}\Delta^{1/2}\sum_{m=-\infty}^{\infty}&\left[J_{2m}(V_0t)\cos\frac{\vartheta}{2}\right.\\&\left.+J_{2m-1}(V_0t)e^{i\varphi}\sin\frac{\vartheta}{2}\right]\\&\times\exp\left[-iKx_0+2imkx_0-\frac{1}{2}\Delta^2(K-2mk)^2\right],\end{aligned}$$

$$\begin{aligned}f_e(K,t)=\pi^{-1/4}\Delta^{1/2}\sum_{m=-\infty}^{\infty}&\left[J_{2m}(V_0t)e^{i\varphi}\sin\frac{\vartheta}{2}\right.\\&\left.+J_{2m+1}(V_0t)\cos\frac{\vartheta}{2}\right]\exp\left\{-iKx_0+i(2m+1)\right.\\&\left.\times kx_0-\frac{1}{2}\Delta^2[K-(2m+1)k]^2\right\}.\end{aligned} \tag{19}$$

For a prolonged interaction of the atom with the field ($V_0t\gg1$), noting the asymptotic behavior of the Bessel function, we conclude that the momenta of the scattered atoms lie within the interval $(-\hbar kV_0t,+\hbar kV_0t)$.

It is important that, even for a broad atomic beam, after averaging over x_0 the momentum distributions of the atoms in states $|g\rangle$ and $|e\rangle$ are asymmetric:

$$\begin{aligned}\overline{|f_g(K,t)|^2}=\frac{\Delta}{2\sqrt{\pi}}\sum_{m=-\infty}^{\infty}&[J_{2m}^2(V_0t)(1+\cos\vartheta)\\&+2J_{2m}(V_0t)J_{2m-1}(V_0t)\\&\times\sin\vartheta\cos\varphi+J_{2m-1}^2(V_0t)(1-\cos\vartheta)]\\&\times\exp[-\Delta^2(K-2mk)^2],\end{aligned}$$

$$\begin{aligned}\overline{|f_e(K,t)|^2}=\frac{\Delta}{2\sqrt{\pi}}\sum_{m=-\infty}^{\infty}&[J_{2m+1}^2(V_0t)(1+\cos\vartheta)\\&+2J_{2m+1}(V_0t)J_{2m}(V_0t)\sin\vartheta\\&\times\cos\varphi+J_{2m}^2(V_0t)(1-\cos\vartheta)]\\&\times\exp\{-\Delta^2[K-(2m+1)k]^2\}.\end{aligned} \tag{20}$$

Figure 1 shows the distribution function averaged over x_0, $\overline{|f_g(K,t)|^2+|f_e(K,t)|^2}$. An asymmetry shows up only when the atomic state is a superposition of the ground and excited states ($\vartheta\neq n\pi$) and is greatest when their contributions are equal ($\vartheta=n\pi+\pi/2$, where n is an integer).

For the integral characteristics of atomic scattering in a standing wave field, it is natural to use the average momentum $\langle p\rangle$ transferred to an atom and its dispersion $p_d=\sqrt{\langle p^2\rangle-\langle p\rangle^2}$, which can be calculated if the average value of the square of the momentum transfer to an atom is known. Here it is more convenient to use a coordinate representation.

The average momentum of an atom is

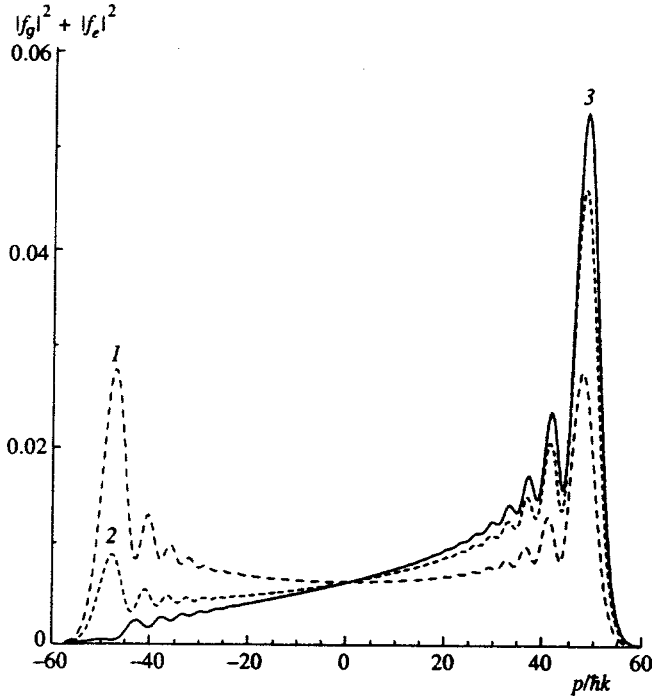


FIG. 1. Momentum distribution function of atoms after scattering by the field of a standing wave in the case in which the initial state of the atom is a superposition of the ground and excited states created by a traveling wave for $\delta=0$. The width of the wave packet of the atom is $\Delta=1/2k$, $V_0t=50$. Curves *1*, *2*, and *3* refer to the cases $\vartheta=0$, $\vartheta=0.25\pi$, and $\vartheta=0.5\pi$.

$$\langle p\rangle=\frac{\hbar k}{2}\{1-\cos\vartheta+V_0t\sin\vartheta[\cos\varphi+\exp(-k^2\Delta^2)\cos(2kx_0+\varphi)]\}. \tag{21}$$

The time-independent part of Eq. (21) describes the momentum acquired by an atom from the traveling wave, and the part which is proportional to the time is due to interaction with the standing wave. In accordance with the preceding remarks, Eq. (21) implies that in the standing wave field, the momentum of an atom varies only when the atom is in a coherent superposition of the ground and excited states at the beginning of the interaction with the standing wave ($\vartheta\neq n\pi$, where n is an integer). The momentum transferred to the atom is greatest for $\varphi=0$, i.e., when the traveling wave is in phase with the one of the traveling waves forming the standing wave that corresponds to it.

The average of the square of the momentum of the atom is

$$\begin{aligned}\langle p^2\rangle&=-\hbar^2\int_{-\infty}^{\infty}\Psi^*(x,t)\,\frac{\partial^2}{\partial x^2}\,\Psi(x,t)dx\\&=\frac{3\hbar^2}{2\Delta^2}+\frac{(\hbar k)^2}{2}\{(V_0t)^2[1+\exp(-k^2\Delta^2)\\&\quad\times\cos(2kx_0)]+2\exp(-k^2\Delta^2)V_0t\sin\vartheta\\&\quad\times\cos(2kx_0+\varphi)+1-\cos\vartheta\}.\end{aligned} \tag{22}$$

As an example, let us consider the case $\vartheta=0.5\pi$, $\varphi=0$, and $kx_0=n\pi$ (atoms interacting with a standing wave near its nodes), when the momentum transfer to the atom is greatest. If the width of the wave packet is much less than the wavelength ($\Delta k\ll 1$), then $p_d\ll\langle p\rangle$ and it can be assumed that the momentum of essentially all atoms changes by the same amount after crossing the standing wave. Note that, according to Eq. (16), in this case the atom is in the diabatic state[12]

$$|\phi_+\rangle=\frac{\sqrt{2}}{2}(|g\rangle+|e\rangle), \tag{23}$$

corresponding to the potential $\hbar V_0\sin(kx)$. It is acted on by a force $F=\hbar kV_0$ which causes the momentum to change by $\hbar kV_0t$ over the time t the standing wave acts.

In the general case of an arbitrary initial position of the atom, its wave function is already a superposition of two diabatic states $|\phi_+\rangle$ and $|\phi_-\rangle$, and the wave packet splits into two, each of which moves in its potential (the optical analog of the Stern–Gerlach experiment). If an atomic beam much wider than the wave length interacts with the field, then, as is easily seen from Eqs. (21) and (22) after integrating over x_0, the average momentum transferred to a beam atom and the change its squared average are independent of the width Δ of the atomic wave packets. This is consistent with the fact that observations of the scattering of a broad atomic beam cannot provide information on the degree of coherence of the plane waves describing an atomic ensemble.[7] Thus, the values observed during scattering of an atomic beam can be obtained by solving the plane wave scattering problem and then averaging the result over the angular distribution of the plane waves. Neglecting the obvious effect of the initial momentum distribution of the atoms on the average momentum and dispersion of the scattered beam, we model it with a plane wave in the following.

We now consider the effect of the offset of the field frequency from that of the atomic transition on the average momentum transferred to an atom and the average square of the momentum. We are interested in the case of a long interaction between the atom and the field, $V_0t\gg 1$. Retaining only linear terms in t in calculating $\overline{\langle p\rangle}$ and quadratic terms in t in calculating $\overline{\langle p^2\rangle}$, we find

$$\overline{\langle p\rangle}=\hbar kV_0^4t\,\frac{\tilde{V}\cos\varphi\sin\vartheta-\delta\sin\varphi(1-\cos\vartheta)}{\tilde{V}^2(2V_0^2+\delta^2+|\delta|\sqrt{4V_0^2+\delta^2})}, \tag{24}$$

and

$$\overline{\langle p^2\rangle}=\frac{(\hbar kV_0^2t)^2}{2V_0^2+\delta^2+|\delta|\sqrt{4V_0^2+\delta^2}}. \tag{25}$$

As can be seen from Eqs. (24) and (25), the average momentum transfer falls off much faster as the offset is increased ($\propto\delta^{-3}$) than does its dispersion ($\propto\delta^{-1}$), and, ultimately, for $\delta\gg V_0$ the momentum distribution of the atoms becomes almost symmetric. The change in the sign of $\overline{\langle p\rangle}$ as δ increases is related mainly to the variation in ϑ and, in the limit $\delta\gg V_0$, takes place with a step size $2\pi/\tau$ determined by the time of the interaction with the travelling wave.

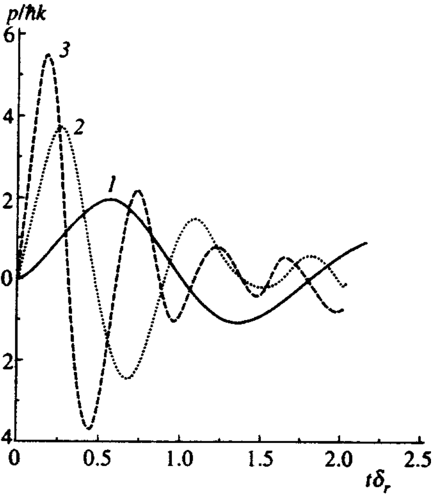


FIG. 2. Average momentum of an atom in the field of counterpropagating rectangular pulses of duration $T=2/\delta_\tau$ as a function of time for $V_0=10\delta_\tau$ (curve *1*), $V_0=40\ \delta_\tau$ (curve *2*), and $V_0=90\delta_r$ (curve *3*). The time shift between the pulses, τ, obeys the condition $V_0\tau=0.5\pi$, where V_0 is the Rabi frequency in the field of a single travelling wave.

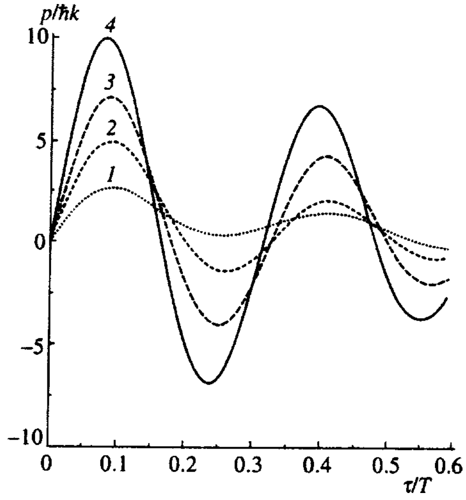


FIG. 3. The average momentum transferred to at atom in the field of counterpropagating gaussian (curve *1*), supergaussian with $s=2$ and $s=4$ (curves *2* and *3*), and rectangular (curve *4*) light pulses as functions of the delay between pulses. For all these curves $V_0\tau=20$, $\delta=0$, and $\delta_\tau=0$.

3. In the above theory we have left the kinetic energy operator, which accounts for the motion of the atom, out of the Hamiltonian H_0. In the general case of an arbitrary duration of the interaction between the atom and the light, initially ($t<T_{RN}$) one should expect a rise in the momentum transfer to the atom with time; then the Doppler effect causes an offset of the atomic transition frequency from the light frequency which should lower the efficiency of the atom-field interaction and limit the further rise in the momentum of the atom.

Noting that for $t<T_{RN}$ the change in the momentum of the atom can reach $\hbar kV_0t$, we may expect that the average momentum transfer to the atom would be of order $\hbar kV_0T_{RN}$.

From the Schrödinger equation (1) with the free-atom Hamiltonian (2), including the translational motion and a field of the form

$$\mathbf{E}=\mathbf{e}[E_0^{(+)}(t)\sin(kx-\omega t)+E_0^{(-)}(t)\sin(kx+\omega t)],\quad(26)$$

after transforming to the momentum representation (17), we can obtain the equations

$$\begin{aligned}\dot f_g(K,t)&=-\frac{i\hbar}{2m}K^2f_g(K,t)+\frac{1}{2}e^{i\delta t}[V_0^{(-)}(t)f_e(K-k,t)\\&\quad-V_0^{(+)}(t)f_e(K+k,t)],\\ \dot f_e(K,t)&=-\frac{i\hbar}{2m}K^2f_e(K,t)+\frac{1}{2}e^{-i\delta t}[V_0^{(+)}(t)f_g(K\\&\quad-k,t)-V_0^{(-)}(t)f_g(K+k,t)],\end{aligned}\quad(27)$$

where $V_0^{(\pm)}(t)=\langle e|\hat{\mathbf{d}}\cdot\mathbf{e}|g\rangle E_0^{(\pm)}(t)/\hbar$; the momentum distribution of the atoms can be found from Eq. (27). Figure 2 shows a set of plots of the average atomic momentum as a function of time obtained by solving Eq. (27) numerically. The calculation has been done for the case of an atom interacting with two counterpropagating waves of equal intensity, that are resonant with the atomic transition and have rectangular envelopes of duration T, amplitude V_0, and a time shift τ between them which has been chosen so that $V_0\tau=0.5\pi$. The values of V_0 corresponding to the different curves are in the proportion $1:2^2:3^2$. It is clear that the corresponding maximum average momenta of the atoms are in the proportion 1:2:3. On the other hand, the times (on the order of T_{RN}) over which the momentum increases to these values are in inverse proportion.

4. We have examined the interaction of an atom with counterpropagating waves in the case where the time dependence of the fields acting on the atom is given by rectangular pulses. A coherent superposition of the ground and excited states is prepared while only the first wave is acting the atom, and then the atom is scattered in a standing wave. Under real conditions, these two processes cannot always be considered independently, so that the formation of a coherent superposition of the ground and excited states prior to scattering in the field of counterpropagating waves with comparable amplitudes and the scattering process, itself, cannot be separated in the general case. One might suppose, for example, that this situation is probable during the interaction of an atomic beam with two spatially shifted counterpropagating laser beams that have gaussian radial intensity distributions.

Figure 3 shows plots of the momentum transferred to an atom beam as a function of the delay between counterpropagating waves for a rectangular envelope, as well as for gaussian and supergaussian light pulses defined by

$$E_0^{(\pm)}=E_0\exp\left[-\left(\frac{2(t\pm\tau/2)}{T}\right)^{2s}\right],\quad(28)$$

found by numerical integration of the Schrödinger equation.

In order to separate effects associated with the motion of the atoms along the laser beam from the effect of the shape of the envelope of the light pulses, Eq. (27) was solved in the Raman–Nath approximation ($t<T_{RN}$). As can be seen from

the figure, s has almost no effect on the delay at which the momentum transferred to the atom beam is greatest, but does have a substantial effect on the magnitude of this momentum. For observing an asymmetry in the scattering of atoms in a standing wave, it is preferable to form this wave from the field of counterpropagating pulses whose envelopes have a near rectangular shape (large s), which essentially excludes the intermediate interaction region where the formation of a coherent superposition of the ground and excited states of the atom can take place simultaneously with scattering of the atom in the field of the counterpropagating waves.

5. We have examined the scattering of atoms in the field of counterpropagating waves and shown that the initial interaction, within a certain time, of an atom with the field of only one of the waves causes the formation of a superposition of the ground and excited states of the atom, whose interaction with the field of both counterpropagating waves differs qualitatively from the scattering of a ground or excited state atom by a standing wave field. The resulting asymmetry in the atomic scattering is caused by a spatial modulation in the amplitude of the excited state of the atom by the field of a traveling wave with a period equal to the wavelength.

This work was supported by the State Foundation for Basic Research of the Ukrainian Ministry of Science and Technology (Grant 2.4/179).

[1] R. J. Cook and A. F. Bernhardt, Phys. Rev. A **18**, 2533 (1978).
[2] A. F. Bernhardt and B. W. Short, Phys. Rev. A **23**, 1290 (1981).
[3] E. V. Baklanov and V. G. Minogin, Zh. Éksp. Teor. Fiz. **92**, 417 (1987) [Sov. Phys. JETP **65**, 237 (1987)].
[4] P. A. Ruprecht, M. J. Holland, and K. Burnett, Phys. Rev. **49**, 4726 (1994).
[5] U. Janicke and M. Wilkens, Phys. Rev. **50**, 3265 (1994).
[6] C. J. Lee, Phys. Rev. **53**, 4238 (1996).
[7] C. S. Adams, M. Siegel, and J. Mlynek, Phys. Rep. **240(3)**, 143 (1994).
[8] V. I. Romanenko and L. P. Yatsenko, JETP Lett. **63**, 968 (1996).
[9] V. A. Grinchuk, E. F. Kuzin, M. L. Nagaeva *et al.*, JETP Lett. **57**, 548 (1993).
[10] G. A. Ryabenko, V. A. Grinchuk, L. A. Grishina *et al.*, Laser Phys. **6**, 150 (1996).
[11] V. G. Minogin and V. S. Letokhov, *The Pressure of Laser Radiation on Atoms* [in Russian], Nauka, Moscow (1986).
[12] A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *The Mechanical Effect of Light on Atoms* [in Russian], Nauka, Moscow (1991).

Translated by D. H. McNeill

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Rotation and alignment of diatomic molecules and their molecular ions in strong laser fields

M. E. Sukharev

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia

V. P. Kraĭnov

Moscow Institute of Physics and Technology, 141700 Dolgoprudnyi, Moscow Region, Russia

(Submitted 6 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 573–582 (February 1998)

A theory of classical rotation and alignment of diatomic molecules with and without permanent dipole moments and of their molecular ions in strong laser fields is developed. The actions of both cw and pulsed laser fields is considered. Conditions under which molecular axes are aligned with the field, which is presumed to be linearly polarized, have been determined. The analysis leads to a conclusion that molecules exposed to ultrashort laser pulses continue to rotate even after the end of the laser pulse. The effect of dynamical chaos on the rotational angular velocity in strong laser field is discussed. *© 1998 American Institute of Physics.* [S1063-7761(98)01402-4]

1. INTRODUCTION

The problem of rotation and alignment of molecules and molecular ions in strong ac field came up quite recently in the analysis of experiments on ionization and dissociation of molecules by intense laser fields.[1–3] Of course, the alignment of molecules is random in the absence of external field. The theoretical problem is in determination of classical dynamics of molecular rotation and oscillation, and also of their orientation in the external ac field. This classical problem is essentially nonlinear because the molecules can rotate. Its solution depends on the pecularities of the molecular system and the laser radiation parameters.

Most experiments with molecules are done at ambient temperature. The characteristic energy of thermal rotation of the diatomic molecule is $kT \simeq 0.025$ eV, where k is Boltzmann's constant and T is the absolute temperature. Even if the laser field intensity is relatively low, $F=5\times 10^7$ V/cm, the typical energy perturbation in a molecule is of the order of $Fa_B \sim 0.27$ eV, i.e., it is considerably higher than the thermal energy, and the latter can be neglected. In this case, the molecule axis should rotate in the plane containing the initial orientation of this axis and the polarization vector of the laser field (in what follows we assume that the light is linearly polarized), so the problem is one-dimensional and there is no precession of the molecular axis caused by the field. In the case of a molecule with zero dipole moment, the characteristic perturbation energy is approximately [see Eq. (1) below] $\alpha F^2/4 \sim 0.03$ eV when the field intensity has the value given above (the molecule's polarizability is set at 30 a.u.; see the example of Cl_2 in Sec. 6). This is comparable to the thermal energy, but this estimate refers to the case of a weak field. At higher field intensities thermal rotation can be neglected.

The internuclear distance in molecules and their ions will be assumed to be fixed, i.e., in our calculations we will ignore the possibility of molecules dissociating while they undergo classical rotation in space. It is clear that the approximation of a frozen internuclear distance not apply to all cases. It presupposes that it is easier to turn a molecule than to stretch it. In the weak-field case this approximation is based on the smallness of rotational energies as compared to vibrational. In the case of a strong field, the approximation is based on an assumption that the typical times of molecular rotations are much shorter than the characteristic times of dissociation. The dissociation rate, in its turn, is slow because is cannot be done without a Landau–Zehner transition from the ground-state electronic level to the separation level. The probability of this transition is exponentially small if the field is below the atomic field, or the internuclear distance is relatively small so that the gap between these electron levels is not very small.

The first parts of the paper are dedicated to the effect of monochromatic laser field with a constant amplitude. The remaining calculations deal with laser pulses of finite amplitude.

2. MOLECULES WITH ZERO DIPOLE MOMENT

Consider a diatomic molecule with zero dipole moment (for example, a molecule of two identical atoms) in a laser field. The angle between the molecule's axis and field polarization vector is denoted by θ. The potential energy of such a system is

$$U=-\frac{1}{2}\alpha F^2 \cos^2 \omega t \cos^2 \theta, \tag{1}$$

where α is the static dipole polarizability of the molecule, F is the laser field amplitude, and ω is its frequency. This formula is derived in the second order of the perturbation theory with respect to the applied field. From the formal viewpoint, the polarizability at the external field frequency ω should be involved, but this frequency is almost always lower than the frequency of electron transitions contributing to a diagonal two-photon element of the polarizability matrix, so the static polarizability can be used. As for the time dependence in Eq. (1), it is meaningful only in cases of a low-frequency field, when the field changes slowly (the criterion is the same as for the field frequency, i.e., the frequency should be low in comparison with the frequency of electron transitions in the molecular system). In the case of an isolated atom, this time dependence is assumed to be in the Stark shift of the energy when the atom is exposed to a laser pulse and the pulse envelope must be taken into account. Thus, it is unnecessary to average Eq. (1) over time. At the same time, the ratio between the frequency ω and the molecule axis rotation frequency can be arbitrary [see Eq. (5) below]. Hence, Newton's equation in this case has the form

$$\frac{d^2\theta}{d\tau^2}+\frac{1}{2}\left(\frac{\omega_0}{\omega}\right)^2\cos^2\tau\sin 2\theta=0. \tag{2}$$

Here the dimensionless time is $\tau\equiv\omega t$ and

$$\omega_0^2=\frac{\alpha F^2}{\mu R^2}, \tag{3}$$

where μ is the reduced mass of the molecule and R is the internuclear distance in the molecule. It is obvious that the frequency ω_0 is the natural frequency of small oscillations of the molecule in a dc electric field. This frequency is undoubtedly much smaller than the frequency of electron transitions in the molecule. After linearization, Eq. (2) turns to a Mathieu equation. It will be shown in the following sections that such a linearization is possible only in weak laser fields.

Equation (2) was solved numerically taking the initial conditions in the form

$$\theta(0)=\frac{\pi}{4},\quad \frac{d\theta}{dt}(t=0)=0. \tag{4}$$

The first of these conditions defines the average angle between the molecular axis and field polarization vector, but in reality this condition has little effect on the solution. The second condition in Eq. (4) derives from the absence of thermal rotations in the molecule, which was discussed above.

Figure 1 shows the function $\theta(\tau)$ for the case of a weak laser field, when $(\omega_0/\omega)^2=10^{-3}$. We can see that the molecular axis oscillates around the laser field polarization vector with the amplitude equal to the initial angle. The oscillation period can be easily calculated analytically using the Krylov–Bogolyubov–Mitropolskii approximation.[4] By linearizing Eq. (2) and replacing the square of the sine with 1/2, we obtain the dimensionless oscillation period

$$T=2\pi\sqrt{2}\,\frac{\omega}{\omega_0}. \tag{5}$$

In this specific case we have 280, which is in a fair agreement with the numerical calculations plotted in Fig. 1.

Figure 2 shows a solution of Eq. (2) with initial conditions (4) for the case of $\omega=\omega_0$. In this case the molecule rotates swiftly and randomly. Sometimes the rotation direction is reversed. This is a typical case of the onset of dynamical chaos.[5] If the chaotic motion is averaged over a large time interval, one can observe a diffusion-like time evolution of the rotation angle. Note that dynamical chaos usually occurs when the applied field frequency is comparable to the natural frequencies of the nonlinear system, which is the case in this example.

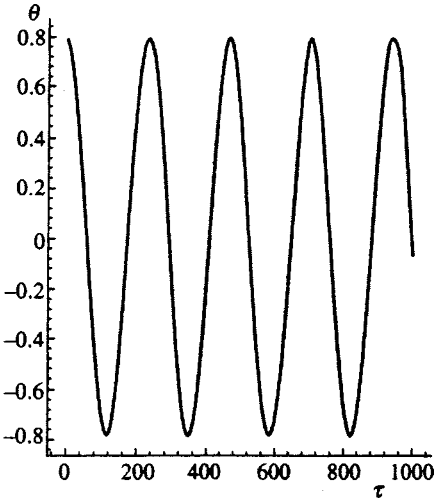


FIG. 1. Deflection angle of a molecule (in radians) as a function of dimensionless time for $(\omega_0/\omega)^2=10^{-3}$.

3. MOLECULAR HYDROGEN ION

Ions of arbitrary molecules are rather difficult to treat when the internuclear distances are small because of the absence of data about the constant dipole moment. In the case of the hydrogen molecular ion the easiest case to study is the configuration in which the internuclear distance R is larger than the equilibrium value. By using a system of units in which the electron charge is unity, we can express the ion potential energy in a monochromatic laser field in the form

$$U=-\frac{1}{2}FR\cos\omega t\cos\theta. \tag{6}$$

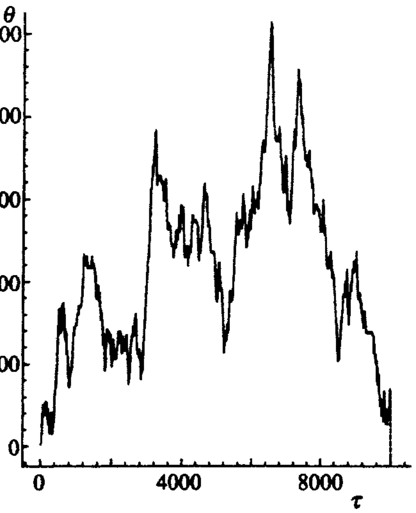


FIG. 2. Same as in Fig. 1, but for $\omega_0=\omega$.

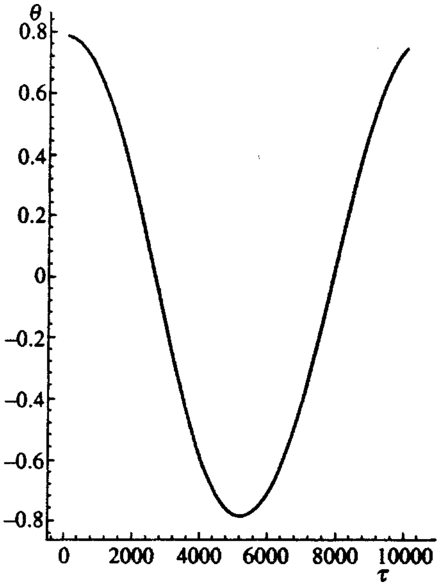


FIG. 3. Deflection angle of the axis of a molecular hydrogen ion (in radians) versus dimensionless time for $(\Omega/\omega)^2=10^{-3}$, $R\gg R_e$.

The notation in this case is identical to that of the previous section. The laser field is assumed to be linearly polarized as previously.

Newton's equation for the angle θ through which the axis of the hydrogen molecule rotates as a function of dimensionless time $\tau=\omega t$ takes the form

$$\frac{d^2\theta}{d\tau^2}+\left(\frac{\Omega}{\omega}\right)^2\cos\tau\sin\theta=0. \tag{7}$$

Here we have introduced the parameter

$$\Omega=\sqrt{\frac{F}{2\mu R}}. \tag{8}$$

This is the frequency of free oscillations of the molecular axis in a dc electric field [similarly to Eq. (3)].

Equation (7) was solved numerically with the initial conditions (4). Figure 3 shows the axis rotation angle as a function of dimensionless time for the case of weak field, when $(\Omega/\omega)^2=10^{-3}$. The curve in Fig. 3 looks like that in Fig. 1. One can see, however, that the period of molecular axis oscillations around the laser field polarization vector is considerably larger than in Fig. 1. This period can be also obtained analytically. By linearizing Eq. (7), we obtain the Mathieu equation with a small parameter:

$$\frac{d^2\theta}{d\tau^2}+\left(\frac{\Omega}{\omega}\right)^2\cos\tau\cdot\theta=0. \tag{9}$$

Its approximate solution can be found using the Whittaker technique.[4] Let us find a solution of a more general Mathieu equation

$$\frac{d^2\theta}{d\tau^2}+\left(\delta+\left(\frac{\Omega}{\omega}\right)^2\cos\tau\right)\theta=0$$

in the form

$$\theta(t)=\theta_0(t)+\left(\frac{\Omega}{\omega}\right)^2\theta_1(t)+\dots$$

around $\delta=0$. The parameter δ should be also expanded in powers of the small parameter. This procedure is in fact equivalent to separation of equations for the fast and slow components of the deflection angle $\theta(t)$.

The coefficients of these power expansions are determined from the condition that secular (divergent) terms not appear in the iteration procedure. In what follows we will construct a solution satisfying the initial conditions (4) from the two linearly independent solutions obtained earlier. After calculations we have

$$\theta(\tau)=\theta(0)\cos\left(\frac{1}{\sqrt{2}}\left(\frac{\Omega}{\omega}\right)^2\tau\right). \tag{10}$$

Thus, the dimensionless oscillation period

$$T=2\pi\sqrt{2}\left(\frac{\omega}{\Omega}\right)^2. \tag{11}$$

In this specific case it is about 9000, which is in good agreement with the numerical solution illustrated by Fig. 3. Note that fast oscillations of the deflection angle with the frequency of the external field, whose amplitude is small, cannot be seen in the graph, given the scale on which is curve is plotted.

Equation (7) was also solved in the case of moderate fields with $\Omega=\omega$. It was found that the molecular ion rotates rapidly and the rotation direction is determined by initial conditions. In the case of strong field, $\Omega\gg\omega$, small fluctuations due to the different phase differences between the field and the rotation angles at different times are superposed on even faster rotations.

The solutions described above correspond to large internuclear separations in a hydrogen molecular ion. At smaller distances, including $R=2$ a.u. the potential energy has a very complex form. It was given in an earlier publication.[6] Using this energy and assuming the laser photon energy to be 0.1 eV (a CO_2 laser), we also solved Newton's equation with initial conditions (4). The solution for the case of the field intensity $F=10^9$ V/cm and internuclear distance $R=4$ a.u. is shown in Fig. 4. One can see that a large-amplitude ion rotation takes place sometimes accompanied by delays. This mode of rotation is also due to specific features of dynamical chaos.

In the case of a weak field and arbitrary internuclear distances, the molecular ion axis slowly oscillates around the laser field polarization vector. This situation is similar to the case illustrated by Fig. 3.

4. MOLECULES WITH A CONSTANT DIPOLE MOMENT

The potential energy of a diatomic molecule with a constant dipole moment d in a laser field is expressed as

$$U=-\frac{1}{2}\alpha F^2\cos^2\omega t\cos^2\theta-dF\cos\omega t\cos\theta. \tag{12}$$

The notation is identical to that of the previous sections.

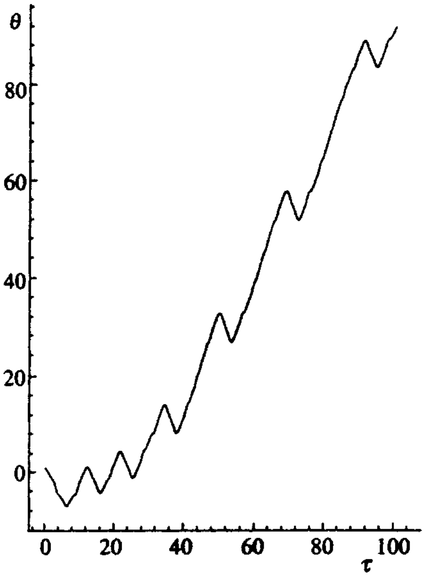


FIG. 4. Same as in Fig. 3, but for $F=10^9$ V/cm, $R=4$ a.u.

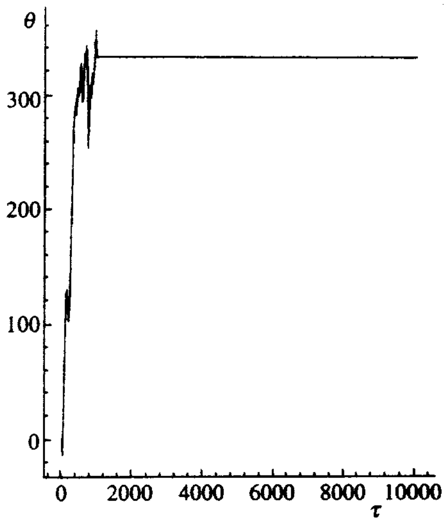


FIG. 5. Rotation angle of a molecule with a constant dipole moment versus dimensionless time for $\Omega=\omega_0=\omega$.

In order to obtain a solution distinct from the two discussed in the previous sections, we need the two terms on the right-hand side of Eq. (12) of the same order of magnitude, namely $F\sim d/\alpha$. For example, in the case of the NO molecule the field intensity amplitude should be $F\sim 2\times 10^7$ V/cm. This field is smaller than the atomic field, therefore the contributions from higher-order hyperpolarizabilities to Eq. (12) can be neglected.

In this case Newton's equation for the molecule rotation angle has the form

$$\frac{d^2\theta}{d\tau^2}+\left(\frac{\Omega}{\omega}\right)^2 \cos\tau \sin\theta+\frac{1}{2}\left(\frac{\omega_0}{\omega}\right)^2 \cos^2\tau \sin 2\theta=0, \tag{13}$$

where the frequency ω_0 is defined by Eq. (3) and

$$\Omega=\sqrt{\frac{dF}{\mu R^2}}. \tag{14}$$

Integration of Eq. (13) with initial conditions (4) for $\omega_0=\Omega=\omega$ yields the function shown in Fig. 5. Unlike the previous cases, the following new effect occurs: the molecule rotates for some time, and after 106 turns it is aligned exactly with the laser field polarization vector. This result derives from the KAM-theorem[5] concerning the stabilization of classical motion in an external field with an amplitude below a certain critical value.

Equation (13) was also integrated in the case of a weak field. In this situation the solution is similar to that shown in Fig. 1, i.e., the molecular axis oscillates around the laser polarization vector with an amplitude equal to the initial angle. In a strong field, on the contrary, the molecule's axis rotates randomly without limits owing to the dynamical chaos effects.

The Schrödinger equation for this case was solved by Charron *et al.*[7] In our opinion, the classical approach to this problem is quite adequate because an external laser field excites the molecule to states of molecular levels under consideration with high rotational quantum numbers. We did not, however, compare this solution with the numerical one because this comparison required initial conditions defined in the form of a quasiclassical wave packet, and such conditions were not introduced in the quantum problem.[7]

5. DISCUSSION OF EXPERIMENTAL DATA

The dynamics of molecular axis rotation in a strong laser field was investigated in experiments.[3] Let us discuss alignment of the neutral Cl_2 molecule in a laser pulse with a duration of 100 fs and a wavelength of 610 nm. The static polarizability of this molecule is $\alpha=31.1$ a.u., the reduced mass of the molecule is $\mu=17.73$ a.m.u., and the equilibrium internuclear distance is $R=3.6$ a.u.[8] Then we derive from Eq. (2), even for $F=1$ a.u.,

$$\left(\frac{\omega_0}{\omega}\right)^2\approx 0.01. \tag{15}$$

Thus, the situation corresponds to the case illustrated by Fig. 1. Figure 6 shows a numerical solution of Eq. (2) taking into account the Gaussian envelope of a laser pulse with a duration of 100 fs. It is clear that the pulse ends after one period of the molecular axis oscillation, then the molecule rotates with a constant velocity imparted by the laser pulse.

When the pulse duration increases to 2 ps, the picture is essentially the same, but the final rotation velocity is two orders of magnitude lower. As the pulse duration increases further, the final velocity tends to zero, in accordance with the well-known adiabatic theorem.

Now let us turn to the singly ionized chlorine molecule ion Cl_2^+. Let us assume that the internuclear distance is larger than the equilibrium value so that the results for the hydrogen molecular ion discussed above could apply. Figure 7 shows the dynamics of the chlorine molecular ion rotation for the case $R=10$ a.u. The other parameters were set equal to those of the neutral chlorine molecule (the pulse duration was 100 fs). It is clear that the molecule oscillates around the

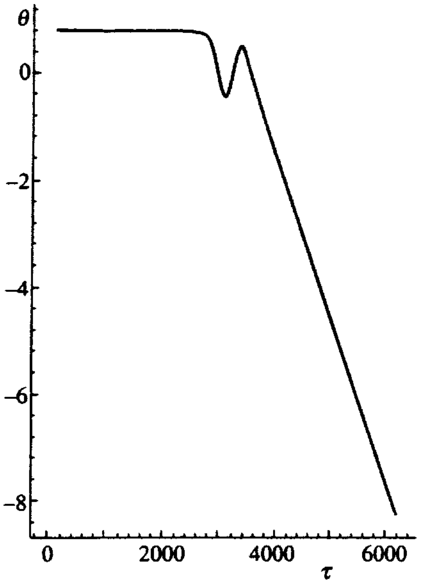


FIG. 6. Rotation angle of Cl_2 molecule versus dimensionless time due to the action of a laser pulse with a duration of 100 fs, a wavelength of 610 nm, and peak intensity of 10^{15} W/cm^2.

polarization vector, and after the end of the laser pulse it rotates at a constant velocity owing to the nonadiabatic property of the pulse envelope.

It is obvious that molecules and their ions do not rotate after a laser pulse if its duration is in the picosecond range, much less if it is in the nanosecond range.

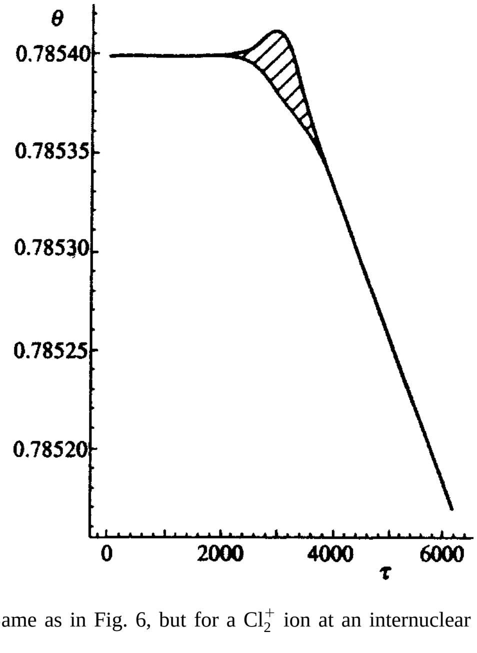


FIG. 7. Same as in Fig. 6, but for a Cl_2^+ ion at an internuclear distance $R = 10$ a.u.

6. CONCLUSIONS

The following conclusions derive from our numerical calculations.

The axes of diatomic molecules with zero dipole moment in a weak laser field oscillate around the polarization vector of the laser radiation (which is assumed to be linearly polarized) with an amplitude equal to the initial angle between the molecule axis and polarization vector. In higher fields effects of dynamic chaos set in, and the molecule rotates randomly. In ultrastrong field the rotation velocity is considerably larger, although irregular features persist in the form of rotation delays at some moments.

The axes of molecular ions in a weak field also oscillate, but at considerably lower frequencies. At higher field intensities they start to rotate. Molecules with constant dipole moments stabilize after fast rotation at certain values of the field intensity when the molecular axis is aligned with the laser field polarization. In ultrastrong fields the molecules rotate rapidly.

Taking into account the shape of an ultrashort laser pulse, we have come to the conclusion that after the pulse end the molecules and their ions rotate at a constant angular velocity determined by initial conditions for the specific molecule. We must stress once again that our conclusions are not affected by specific values of the initial angle between the molecular axis and polarization vector chosen for the numerical calculations.

We are indebted to S. P. Goreslavskii, N. B. Delone, A. M. Popov, and M. V. Fedorov for stimulating discussions of topics concerning this work.

The work was supported by the Russian Fund for Fundamental Research (Grants Nos. 96-02-18299 and 95-02-03657).

[1] F. A. Ilkov, T. D. G. Walsh, S. Turgeon, and S. L. Chin, Phys. Rev. A **51**, R2695 (1995).
[2] S. L. Chin, A. Talebpour, T. D. G. Walsh, S. Larochelle, F. A. Ilkov, and C. Y. Chien, in *Multiphoton Processes 1996*, ed. by P. Lambropoulos and H. Walther, IOP, Bristol (1997), p. 266.
[3] D. Normand, S. Dobosz, M. Lezius, P. D'Oliviera, and M. Schmidt, *ibid*, p. 287.
[4] A. Nayfeh, *Introduction to Perturbation Techniques*, Wiley, New York (1981).
[5] A. Lichtenberg and M. Lieberman, *Regular and Stochastic Motion*, Springer, Heidelberg (1982).
[6] M. E. Sukharev and V. P. Krainov, Laser Phys. **7**, 803 (1997).
[7] E. Charron, A. Giusti-Suzor, and F. Mies, Phys. Rev. A **49**, R641 (1994).
[8] A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions*, Atomizdat, Moscow (1980); Springer, Berlin (1985).

Translation provided by the Russian Editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Ionization of a molecular hydrogen ion by a strong low-frequency electromagnetic field of laser radiation

M. B. Smirnov and V. P. Kraĭnov

Moscow Physicotechnical Institute, 141700 Dolgoprudnyĭ, Moscow Region, Russia

(Submitted 15 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 583–592 (February 1998)

Simple analytical expressions are obtained for the energy and angular distributions of outgoing electrons in ionization of a molecular hydrogen ion by a strong low-frequency electromagnetic field as well as for the ionization probabilities per unit time. The cases of linear and circular polarization of the laser radiation are studied. It is shown that in contrast to the case of the ionization of atoms oscillations appear in the energy spectra of the photoelectrons as a function of their kinetic energy. The well-known limits for the tunneling ionization probabilities for the hydrogen atom by a strong low-frequency alternating field are obtained in the case of large internuclear separations.
[S1063-7761(98)01502-9]

1. INTRODUCTION

A series of theoretical and experimental works on the ionization of simple molecular systems by a strong field of low-frequency laser radiation has appeared in the last few years.[1–5] In this case the ionization is of a tunneling character and there are no resonance effects. Specifically, in Ref. 6 the problem of the one-dimensional dynamics of a molecular hydrogen ion with fixed internuclear separation was solved by numerical methods. Three-dimensional calculations were performed both with a fixed internuclear separation and with internuclear separation varying in the process of dissociation.[7] A useful simplification of the calculations is the two-level approximation, which takes account of only the lowest even and odd states of the molecular hydrogen ion. The problem is also simplified by the fact that the axis of the molecular ion is aligned in the direction of the field within times much shorter than the duration of the laser pulse. This has been confirmed theoretically[8] and experimentally.[5]

In the present paper the Keldysh–Faisel–Reiss approach[9] is used to study analytically the ionization of a molecular hydrogen ion by a low-frequency laser field. The Coulomb correction to the Volkov wave function of the final state in the continuous spectrum is taken into account in a well-known approximation.[10] Both linear and circular polarizations are studied. The energy and angular distributions of the outgoing photoelectrons in the tunneling regime and the ionization probabilities per unit time are found.

2. COULOMB CORRECTION

In the Keldysh–Faisel–Reiss approach[9] the so-called Volkov wave function describing the motion of an electron in the field of only an electromagnetic wave corresponds to the wave function of the final state in the continuous spectrum. The core potential is taken into account in semiclassical perturbation theory. This procedure reduces to the appearance of an additional factor of the form

$$I=\exp\left\{-i\int U dt\right\}. \tag{1}$$

in front of the Volkov wave function. Here U is the potential energy of an electron in the field of the two nuclei of the molecular hydrogen ion

$$U=-\frac{1}{r_a}-\frac{1}{r_b}, \tag{2}$$

which consists of two potential wells. The quantities r_a and r_b are the distances between the electron and the protons a and b, respectively. Here and below atomic units $e=\hbar=m_e=1$ are used everywhere.

We calculate first one of the two integrals in Eq. (1) by replacing the integration over time by integration over the spatial coordinate:

$$\int \frac{dt}{r_b}=\int_{r_b}^{E_i/F}\frac{dr_b}{r_b\sqrt{2(-E_i+Fr_b)}}=-in\,\ln\frac{2}{n^2Fr_b}. \tag{3}$$

Here the integral is calculated in the subbarrier region, $E_i>0$ is the binding energy of the initial electronic level of the molecular hydrogen ion, F is the amplitude of the intensity of the laser radiation field, and $n=1/\sqrt{2E_i}$ is the effective quantum number in the potential well that is closest with respect to the emission of an electron. We assume that the axis of the molecule is directed along the electric field vector, since the ionization probability is much higher in this case than for other orientations of the molecular axis.

The upper limit of integration in Eq. (3) corresponds to the classical turning point $r_b=E_i/F$. Large distances introduce only a phase factor in the Coulomb correction. Only the subbarrier region of distances determines the important contribution.

The contribution from the potential of the distant potential well is calculated similarly

$$\int \frac{dt}{r_a} = -in^* \ln \frac{2}{n^{*2} F r_a}, \tag{4}$$

where the effective principal quantum number for the distant well has the form

$$n^* = \frac{1}{\sqrt{2(E_i + FR)}},$$

where R is the internuclear separation and the quantity FR is the shift between the levels in the left- and right-hand potential wells.

Therefore the Coulomb correction equals

$$I(H_2^+) = \left(\frac{2}{n^{*2} F r_a}\right)^{n^*} \left(\frac{2}{n^2 F r_b}\right)^{n} \geqslant 1. \tag{5}$$

If $R \to \infty$, then $n^* \to 0$, $n \to 1$ and, as should be the case, the factor (5) goes over the well-known Coulomb correction for the hydrogen atom[10]

$$I(H) = \frac{2}{Fr}. \tag{6}$$

In the opposite limiting case $R \to 0$ we obtain from Eq. (5) the Coulomb correction for the hydrogen-like ion of the helium atom.

3. MATRIX ELEMENT

The amplitude of the transition from the initial state i of the molecular hydrogen ion into the final state of the continuous spectrum with electron momentum $\mathbf{p}$ has the following form in the Keldysh–Faisel–Reiss model[11] taking account of the Coulomb correction:[12]

$$A_{i\mathbf{p}} = -i \int \langle \Psi_{\mathbf{p}}^{(V)} I | \hat{V} | \Psi_i^{(0)} \rangle dt. \tag{7}$$

Here the Coulomb correction I is given by the relation (5), and the final state in the continuous spectrum is described by the Volkov wave function

$$\Psi_{\mathbf{p}}^{(V)} = \exp\left\{ i\mathbf{pr} - \frac{i}{2} \int^t \left(\mathbf{p} + \frac{1}{c}\mathbf{A}\right)^2 dt' \right\}, \tag{8}$$

where $\mathbf{A}$ is the vector potential of the electromagnetic field and

$$\hat{V} = \frac{1}{c}\mathbf{pA} + \frac{1}{2}\mathbf{A}^2 \tag{9}$$

is the interaction potential between the electron and the electromagnetic field. Finally, $\Psi_i^{(0)}$ is the unperturbed wave function of the initial electronic state of the molecular hydrogen ion with binding energy E_i.

Integrating by parts following Ref. 13 simplifies the expression for the transition amplitude (7)

$$A_{i\mathbf{p}} = \left(\frac{1}{2}p^2 + E_i\right) \int \langle \Psi_{\mathbf{p}}^{(V)} | I | \Psi_i^{(0)} \rangle dt. \tag{10}$$

In this section we shall calculate the matrix element appearing in the expression (10) for the transition amplitude. For this, we study first the form of the unperturbed wave function of the initial electronic state of the molecular hydrogen ion. There are two close states: the even ground state g and the odd first excited state u. They are strongly mixed by the constant part of the external electromagnetic field as a result of the approximate degeneracy. Therefore

$$\Psi_i^{(0)} = (\cos\beta \cdot \varphi_g + \sin\beta \cdot \varphi_u)\exp(iE_g t). \tag{11}$$

The mixing angle β and the perturbed energy E_g of the ground state can be easily found in the two-level approximation.[14]

Using the variational approximation for the hydrogen orbitals, we write the unperturbed wave functions of the even and odd states in the form[15]

$$\varphi_g = \frac{\alpha^{3/2}}{\sqrt{2\pi(1+S)}} \{\exp(-\alpha r_a) + \exp(-\alpha r_b)\}. \tag{12}$$

and

$$\varphi_u = \frac{\alpha^{3/2}}{\sqrt{2\pi(1-S)}} \{\exp(-\alpha r_a) - \exp(-\alpha r_b)\}. \tag{13}$$

Here α is a variational parameter ($a \approx 1$ for $R \geqslant 3-4$ a.u.) and S is the overlap integral of the orbitals, i.e.

$$S = \langle \exp(-\alpha r_a) | \exp(-\alpha r_b) \rangle.$$

We note that $S \to 0$ for $R \gg 1$.

The variational wave function (13) of the odd state is unsatisfactory for small internuclear separations. In reality, however, we shall not require the internuclear separations to be small, since in the process of dissociation of a molecular hydrogen ion they are greater than the equilibrium internuclear separation, equal to 2 a.u.

So we must actually calculate integrals of the form

$$K_{a,b} = \int \exp\{-\alpha r_{a,b} - i\mathbf{pr}\} I d\mathbf{r}. \tag{14}$$

This is most simply done in elliptic coordinates

$$r_a = \frac{1}{2}R(\xi + \eta), \quad r_b = \frac{1}{2}R(\xi - \eta),$$
$$1 < \xi < \infty, \quad -1 < \eta < 1. \tag{15}$$

The volume element is given by the relation

$$d\mathbf{r} = \pi R r_a r_b d\xi\, d\eta.$$

Next we have

$$-i\mathbf{pr} = -i\mathbf{p}(\mathbf{r}_a - \mathbf{R}/2) = -ip(\xi\eta + 1)R/2 + ip_{\parallel}R/2.$$

Here $p_{\parallel}$ is the projection of the electron momentum on the axis of the molecular ion.

In calculating the integral (14) we make a series expansion in the small parameter $\eta/\xi \ll 1$, corresponding to the emission of an electron in directions close to the axis of the ion. Then we obtain

$$K_a = Z \int \int \exp(-i\mathbf{pr} - \alpha r_a)\xi^{\chi} d\xi\, d\eta, \tag{16}$$

where

$$Z=2\pi R\left(\frac{R}{2}\right)^{\chi}\left(\frac{2}{n^2F}\right)^{n}\left(\frac{2}{n^{*2}F}\right)^{n^*} \tag{17}$$

and

$$\chi=2-n-n^*. \tag{18}$$

We find from Eq. (16)

$$K_a=\frac{2\pi}{ip}\,Z\,\exp(ip_{\parallel}R/2)L\left(x,\,\frac{i\alpha}{p}\right). \tag{19}$$

Here we have introduced

$$L(x,y)\equiv\int_1^{\infty}\frac{\exp[-x^*(\xi-1)]-\exp[-x(\xi+1)]}{\xi-y}\,\xi^{\chi}d\xi \tag{20}$$

and

$$x=\frac{1}{2}\,(\alpha+ip)R.$$

Similarly

$$K_b=-\frac{2\pi}{ip}\,Z\,\exp(-ip_{\parallel}R/2)L\left(x^*,-\frac{i\alpha}{p}\right). \tag{21}$$

It is easy to see that the integral

$$\int\,\exp(-i\mathbf{pr})\varphi_g d\mathbf{r}$$

is real, while the integral

$$\int\,\exp(-i\mathbf{pr})\varphi_u d\mathbf{r}$$

is imaginary.

The probability of ionization of the mixed state described by the wave function (11) per unit time is given by the relation

$$w_{i\mathbf{p}}=\cos^2\,\beta\cdot w_{g\mathbf{p}}+\sin^2\,\beta\cdot w_{u\mathbf{p}}. \tag{22}$$

Here w_{gp} is the ionization probability of the even state and w_{up} is the ionization probability of the odd state.

We shall now find a simple, approximate expression for the integral (20), using the inequalities $pR\gg\alpha R\gg1$ which are satisfied for final-state energies of the electron greater than the Rydberg energy. Integrating by parts, we find

$$L\left(x,\,\frac{i\alpha}{p}\right)\approx\frac{2}{ipR}. \tag{23}$$

Then we obtain for the matrix elements the simple expressions

$$\langle\exp(i\mathbf{pr})|I|\varphi_g\rangle=\frac{8\pi\alpha^{3/2}Z}{\sqrt{2\pi(1+S)}}\,\frac{\cos\,p_{\parallel}R}{p^2R}. \tag{24}$$

Similarly

$$\langle\exp(i\mathbf{pr})|I|\varphi_u\rangle=-i\,\frac{8\pi\alpha^{3/2}Z}{\sqrt{2\pi(1-S)}}\,\frac{\sin\,p_{\parallel}R}{p^2R}. \tag{25}$$

4. LINEAR POLARIZATION

The probability of a transition over the time t from an electronic state i of the discrete spectrum of a molecular ion into a state with momentum $\mathbf{p}$ in the continuous spectrum is given by the square of the absolute value of the transition amplitude

$$W_{i\mathbf{p}}=\left(E_i+\frac{1}{2}\,p^2\right)^2|\langle\exp(i\mathbf{pr})|I|\varphi_i\rangle|^2 \times\left|\int_0^t\exp[ig_L(t')dt']\right|^2. \tag{26}$$

Here the function $g_L(t)$ is determined by the relation

$$g_L(t)\equiv\left(E_i+\frac{1}{2}\,p^2+\frac{F^2}{4\omega^2}\right)t+\frac{p_{\parallel}F}{\omega^2}\cos\,\omega t -\frac{F^2}{8\omega^3}\sin\,2\omega t. \tag{27}$$

The projection of the momentum on the direction of polarization of the laser field has the form

$$p_{\parallel}=p\,\cos\,\theta\approx p(1-\theta^2/2).$$

It is assumed that the laser pulse is long compared with the period $T=2\pi/\omega$ of the oscillations.

To calculate the time integral we note that

$$\exp[ig_L(t)]=\exp(-iS_0)\exp[ig_L(t+2\pi/\omega)], \tag{28}$$

where

$$S_0=\frac{2\pi}{\omega}\left[\frac{1}{2}\,p^2+E_i+\frac{F^2}{4\omega^2}\right]. \tag{29}$$

Thus, for the time $t=2\pi K/\omega$ (K is the integral number of periods of the field) we find the following expression for the integral over the time in Eq. (26):

$$J\equiv\left|\int_0^t\exp[g_L(t')]dt'\right|^2 =\left|\int_0^{2\pi/\omega}\exp[g_L(t)]dt\right|^2\frac{\sin^2(KS_0/2)}{\sin^2(S_0/2)}. \tag{30}$$

Making the substitution $S_0\rightarrow S_0-2\pi N$, where N is an integer (the number of absorbed laser photons), we pass to the limit $K\rightarrow\infty$

$$J=\left|\int_0^{2\pi/\omega}\exp[ig_L(t)dt]dt\right|^2\frac{\omega^2t}{2\pi}\,\delta\left\{\frac{1}{2}\,p^2 +E_i+\frac{E^2}{4\omega^2}-N\omega\right\}. \tag{31}$$

Multiplying by the density of final states $d\mathbf{p}/(2\pi)^3$, dividing by the time t, and integrating over the magnitude of the momentum, we obtain the emission probability of an electron per unit time into a given solid angle $d\Omega$ with the absorption of N photons and with a fixed final momentum of the electron:

$$\frac{dw_{i\mathbf{p}}}{d\Omega}=\frac{\omega^2 p}{2\pi}\left(E_i+\frac{1}{2}p^2\right)^2|\langle\exp(i\mathbf{pr})|I|\varphi_i\rangle|^2 \times\left|\int_0^{2\pi/\omega}\exp[ig_L(t)dt]dt\right|^2. \tag{32}$$

The momentum in the final state is determined from the law of conservation of energy with absorption of N photons

$$\frac{1}{2}p^2+E_i+\frac{F^2}{4\omega^2}=N\omega. \tag{33}$$

The integral over one period of the oscillations can be calculated by the saddle-point method assuming a tunneling regime of ionization

$$\gamma=\frac{\omega\sqrt{2E_i}}{F}\ll 1.$$

We obtain

$$\left|\int_0^{2\pi/\omega}\exp[ig_L(t)]dt\right|^2=\frac{\omega^2}{F\sqrt{2E_i}}\exp\left\{-\frac{2(2E_i)^{3/2}}{3F}\times\left(1-\frac{\gamma^2}{10}\right)-\frac{p_\perp^2\sqrt{2E_i}}{F}-\frac{p_\parallel^2\gamma^3}{3\omega}\right\} \tag{34}$$

(see also Ref. 15). Here $p_\perp=p\sin\theta\approx p\theta$ is the projection of the electron momentum on a plane normal to the polarization vector of the laser field. The axis of the molecular ion is assumed to be directed along the polarization vector of the field.

Using the expressions obtained above for the matrix element, we find the energy and angular distributions of the emitted electrons

$$\frac{dw_{i\mathbf{p}}}{d\Omega}=\frac{\omega^4C_i^2}{2(2\pi)^4p^3F\sqrt{2E_i}}\{1+\cos p_\parallel R\} \times\exp\left\{-\frac{2(2E_i)^{3/2}}{3F}\left(1-\frac{\gamma^2}{10}\right)-\frac{p_\perp^2\sqrt{2E_i}}{F}-\frac{p_\parallel^2\gamma^2}{3\omega}\right\}. \tag{35}$$

Here

$$C_i=\frac{16\pi\alpha^{3/2}R^{\chi-1}}{\sqrt{2\pi(1\pm S)}}\left(\frac{E_i+FR}{F}\right)^{n*}\left(\frac{E_i}{F}\right)^n. \tag{36}$$

Depending on the initial (perturbed) state, the energy E_i equals either E_g or E_u (and, correspondingly, the $+$ sign or $-$ sign is chosen in Eq. (36)).

Integrating over the angle of emission of the electron, we obtain the energy spectrum of the outgoing photoelectrons

$$w_{i\mathbf{p}}=\frac{\omega^4C_i^2}{(2\pi)^4p^3F\sqrt{2E_i}}\left(E_i+\frac{1}{2}p^2\right)^2(1+\cos pR) \times\exp\left\{-\frac{2(2E_i)^{3/2}}{3F}\left(1-\frac{\gamma^2}{10}\right)-\frac{p^2\gamma^3}{3\omega}\right\}. \tag{37}$$

In this expression the oscillatory factor, which originates from interference, is what most distinguishes the tunneling ionization of a molecular ion from tunneling ionization of an atom.

Summing over all numbers N of absorbed photons we find the ionization probability per unit time

$$w_i=\sum_N w_{i\mathbf{p}}=\int\frac{1}{\omega}w_{i\mathbf{p}}d(N\omega)=\int_0^\infty w_{i\mathbf{p}}\frac{p}{\omega}dp. \tag{38}$$

Here the law of conservation of energy (33) with absorption of photons is taken into account. The calculation of the integral in Eq. (38), after substituting the expression (37), gives the final result

$$w_i=\frac{\alpha^3(2E_i)^{3/4}}{2(1\pm S)}\sqrt{\frac{3}{\pi F}}\left(\frac{8E_i}{FR}\right)^{2-2\chi}\left(1+\frac{FR}{E_i}\right)^{2n*} \times\exp\left\{-\frac{2(2E_i)^{3/2}}{3F}\left(1-\frac{\gamma^2}{10}\right)\right\}. \tag{39}$$

We note that the oscillatory interference factor vanishes on integration over the energy of the outgoing photoelectron. In the limit $R\to\infty$ the expression (39) passes, as it should, into the well-known expression for the probability of tunneling ionization by a low-frequency linearly polarized field (see, for example, Ref. 16).

5. CIRCULAR POLARIZATION

In the case of a circularly polarized field the function $g(t)$ has the form

$$g_C(t)=\left(E_i+\frac{1}{2}p^2+\frac{F^2}{2\omega^2}\right)t+\frac{pF}{\omega^2}\cos\psi\cos\omega t. \tag{40}$$

Here ψ is the small angle between the direction of emission of the electron and the polarization plane of the laser radiation.

We calculate the integral

$$J=\left|\int_0^t\exp[ig_C(t)]dt\right|^2$$

by the same method as for the case of linear polarization, expounded in detail in the preceding section. We obtain

$$J=2\pi t\sum_{N=N_0}^{\infty}J_N^2\left(\frac{pF\cos\psi}{\omega^2}\right)\delta\left(\frac{1}{2}p^2+E_i+\frac{F^2}{2\omega^2}-N\omega\right). \tag{41}$$

In the tunneling regime (adiabaticity parameter $\gamma\ll 1$) Debye's asymptotic representation of the Bessel function for large argument and large index can be used to calculate the expression (41):

$$J_N\left(\frac{N}{\cosh\beta}\right)\approx\frac{1}{\sqrt{2\pi N\beta}}\exp\{-N(\tanh\beta-\beta)\}. \tag{42}$$

Then we find

$$J=\sum_{N=N_0}^{\infty}\frac{1}{N\gamma}\delta\left\{\frac{1}{2}p^2+E_i+\frac{F^2}{2\omega^2}-N\omega\right\}$$

$$\times\exp\left\{-\frac{2(2E_i)^{3/2}}{3F}\left(1-\frac{\gamma^2}{15}\right)\right\}\exp\left\{-\frac{F\sqrt{2E_i}}{\omega^2}\,\psi^2 - \frac{\omega^4\sqrt{2E_i}}{F^3}\,(\Delta N)^2\right\}. \tag{43}$$

Here

$$\Delta N \equiv N - \frac{F^2}{\omega^3} - \frac{4E_i}{3\omega}.$$

Substituting the expression (43) into Eq. (32), we find the energy and emission-angle distributions of the emitted electrons

$$\frac{dw_{i\mathbf{p}}}{d\Omega} = \frac{\omega C_i^2}{8(2\pi)^3\sqrt{2E_i}}\exp\left\{-\frac{2(2E_i)^{3/2}}{3F}\left(1-\frac{\gamma^2}{15}\right) - \frac{F\sqrt{2E_i}}{\omega^2}\,\psi^2\right\}\sum_{\Delta N}(1+\cos p_{\parallel}R) \times |\langle \exp(i\mathbf{pr})|I|\varphi_i\rangle|^2 \exp\left\{-\frac{\omega^4\sqrt{2E_i}}{F^3}\,(\Delta N)^2\right\}. \tag{44}$$

Here the coefficient C_i is determined by the expression (36). The quantity $p_{\parallel}$ is the projection of the electron momentum on the axis of the molecular ion.

Integrating over the angles and momenta of the outgoing photoelectron we find the ionization probability per unit time

$$w_i = \frac{2E_i\alpha^3}{\pi(1\pm S)}\left(\frac{8E_i}{FR}\right)^{2-2\chi}\left(1+\frac{FR}{E_i}\right)^{2n^*} \times\left\{1+\sqrt{\frac{2\omega}{\pi FR}}\cos\left(\frac{\pi R}{\omega}-\frac{\pi}{4}\right)\right\}. \tag{45}$$

In the limit of large internuclear separations $R\rightarrow\infty$ this expression reduces, as it should, the probability of ionization of a hydrogen atom by a constant electric field[17]

$$w_i = \frac{4}{F}\exp\left(-\frac{2}{3F}\right).$$

In contrast to the case of linear polarization, small oscillations of the ionization probability as a function of the laser intensity remain in Eq. (45).

6. CONCLUSIONS

In this paper we have obtained for the first time analytical expressions for the probability of ionization of a molecular hydrogen ion per unit time by low-frequency linearly and circularly polarized laser radiation fields (in the tunneling regime). The corresponding energy and emission-angle distributions of the outgoing electrons were also obtained. They are functions of the internuclear separation, and for large separations they pass into the well-known expressions for the probability of tunneling ionization of a hydrogen atom by a low-frequency alternating field. For moderate values of the internuclear separation, the ionization probability was found to exhibit an interference dependence on this distance as a result of the coherent summation of the transition amplitudes in the fields of the different nuclei. The oscillatory dependence is observed only in the energy spectrum of the outgoing photoelectrons. After integration over the energies the oscillations vanish completely in the case of a linearly polarized field and remain partially in the case of a circularly polarized field. All calculations were performed for the case when the axis of the molecular ion is oriented along the polarization vector of the linearly polarized radiation and in the polarization plane of the circularly polarized radiation.

We wish to express our deep appreciation to S. P. Goreslavskiĭ, N. B. Delone, A. M. Popov, and M. V. Fedorov for valuable suggestions concerning the content of this paper. The work was supported in part by the ''Soros Student'' program and the Russian Fund for Fundamental Research (Grants Nos. 95-02-03657 and 96-02-18299).

[1] A. Giusti-Suzor, F. H. Mies, L. F. DiMauro, E. Charron, and B. Yang, J. Phys. B: At. Mol. Opt. Phys. **28**, 309 (1995).
[2] S. L. Chin, A. Talebpour, T. D. G. Walsh, S. Larochelle, F. A. Ilkov, and C. Y. Chien in *Multiphoton Processes 1996*, Institute of Physics Conference Ser. No. 154, IOP Publishing Ltd., Bristol, 1997, p. 266.
[3] S. Chelkowski, A. D. Bandrauk, and A. Conjusteau, *ibid.*, p. 192.
[4] P. Dietrich and M. Yu. Ivanov, *ibid.* p. 202.
[5] D. Normand, S. Dobosz, M. Lezius, P. D'Oliveira, and M. Schmidt, *ibid.*, p. 287.
[6] M. E. Sukharev and V. P. Krainov, Laser Phys. **7(3)**, (1997).
[7] Z. Mulyukov, M. Pont, and R. Shakeshaft, Phys. Rev. A **52**, 206 (1996).
[8] M. E. Sukharev and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **113**, 573 (1998) [JETP **86**, 318 (1998)].
[9] H. R. Reiss, Progr. Quant. Electr. **16**, 1 (1992).
[10] V. P. Kraĭnov and B. Shorki, Zh. Éksp. Teor. Fiz. **107**, 1180 (1995) [JETP **80**, 657 (1995)].
[11] N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms*, Springer, Berlin, 1992.
[12] V. P. Krainov, J. Opt. Soc. Am. B **14**, 425 (1997).
[13] H. R. Reiss, Phys. Rev. A **22**, 1786 (1980).
[14] P. Dietrich and P. B. Corkum, J. Chem. Phys. **97**, 3187 (1992).
[15] N. B. Delone and V. P. Krainov, J. Opt. Soc. Am. B **8**, 1207 (1991).
[16] N. B. Delone and V. P. Krainov, *Atoms in Strong Light Fields*, Springer, Berlin, 1985.
[17] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed., Pergamon Press, New York (1977) [Moscow, Nauka, 1989].

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Numerical modeling of the photoionization of Rydberg atoms by the field of an electromagnetic wave

E. A. Volkova, A. M. Popov,*) and O. V. Tikhonova

D. V. Skobel'tsin Scientific-Research Institute of Nuclear Physics, M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

(Submitted 13 August 1997)
Zh. Éksp. Teor. Fiz. **113**, 593–605 (February 1998)

The ionization of excited hydrogen-like atoms in a femtosecond laser pulse is studied by direct numerical integration of the time-dependent Schrödinger equation for a quantum system in the field of an electromagnetic wave. Expressions are obtained for the ionization probability of the system as a function of the parameters of the laser pulse and the initial state of the atom. Ionization suppression is found, confirming the basic premises of the theory of interference stabilization of Rydberg atoms. © *1998 American Institute of Physics.* [S1063-7761(98)01602-3]

1. INTRODUCTION

One of the most interesting physical phenomena in the physics of the interaction of intense light with atomic and molecular systems is ionization suppression in intense electromagnetic fields.[1–3] This effect is manifested as a decrease in the ionization probability of the system with increasing radiation intensity when the intensity of the field of the electromagnetic wave exceeds some critical value or as saturation of the ionization probability at a level less than 1 and observation of a residual atomic population that does not decrease with increasing intensity of the applied pulse.

At the present time two mechanisms of stabilization in the field of an electromagnetic wave are ordinarily distinguished: interference stabilization, first studied in Ref. 4, and stabilization in the Kramers–Henneberger regime.[5]

According to Ref. 4 interference stabilization arises as a consequence of coherent secondary filling of a group of close-lying Rydberg levels of an atom as a result of the coupling of the levels via the continuum by Λ-type Raman transitions and subsequent destructive quantum interference of the amplitudes of the transition into the continuum from these coherently populated states. The threshold for the appearance of this regime is determined by the overlapping of the ionization widths of neighboring Rydberg states and can be easily estimated in accordance wtih Fermi's golden rule. Using WKB expressions, found in Ref. 6, for the matrix elements V_{nE} coupling the states of the discrete spectrum and the continuum, we obtain the following estimate for the threshold intensity of the field of the wave:

$$\varepsilon^* \propto \omega^{5/3}, \tag{1}$$

where ε and ω are the intensity and frequency of the field of the electromagnetic wave (atomic units are used).

The main generalizations and an elaboration of the theory of interference stabilization are contained in Refs. 7–9.

An alternative stabilization mechanism can be realized even for a system with a single state in a discrete spectrum, but it requires, as a rule, superatomic optical fields. In this case it is convenient to use the Kramers noninertial coordinate system, oscillating as a free electron in the field of the electromagnetic wave, to investigate the dynamics of the system. In this coordinate system an electron ''sees'' the atomic nucleus oscillating with the frequency and amplitude of the oscillations of a free electron. The potential of the nucleus averaged over one optical cycle is the Kramers–Henneberger potential, and in strong fields it has the characteristic double-well shape. Under certain conditions the quantum system demonstrates properties (structure of the energy spectrum) which are characteristic for the Kramers–Henneberger potential and is described well in terms of the stationary states of this potential, the states being highly stable with respect to ionization. Numerous investigations[10–15] show the physical reality of the existence of the Kramers–Henneberger potential describing the ''atom+external electromagnetic field'' system and a decrease in the rate of ionization from it in fields above a threshold value.

In recent years the data obtained by direct numerical integration of the time-dependent Schrödinger equation for a quantum system, taking account of the field of the electromagnetic wave, have been playing a large role in the study of the stabilization of atomic systems in intense light fields. The data so obtained can be viewed as being the results of numerical experiments and provide a reliable basis for assessing the effectiveness of different analytical approaches.

For example, in Ref. 16 the three-dimensional time-dependent Schrödinger equation for a hydrogen atom in the 1s ground state in a linearly polarized field of an electromagnetic wave was solved numerically for the parameters $\hbar\omega = 27.2$ eV and $P = 4\cdot 10^{17} - 6.4\cdot 10^{18}$ W/cm^2. The results confirmed the fact that the rate of ionization of the system decreases with increasing radiation intensity. The dichotomous structure of the electron wave function found in these calculations is an important feature of the manifestation of the wave functions of the stationary states of the Kramers–Henneberger potential and proves that the Kramers–Henneberger method is effective for describing stabilization in the indicated range of the parameters.

We also mention Ref. 17, where a numerical solution was obtained for the three-dimensional Schrödinger equation describing the dynamics of ionization of a hydrogen atom from the $2p$ state for different values of the azimuthal quantum number m by a circularly polarized electromagnetic field. It was shown that the stability of the system against ionization is determined by the quantum number m, and it was conjectured that the effect can be explained on the basis of the interference mechanism.[4]

At present there are few experimental works on the stabilization phenomenon. The investigation of stabilization in the Kramers–Henneberger regime exceeds the capabilities of modern laser technology. As regards interference stabilization, some data attesting to an increase in the ionization stability of Rydberg atoms with increasing radiation intensity were obtained in Refs. 18–20.

Our objective in the present work is to investigate the ionization of a real three-dimensional hydrogen-like atom in the range of parameters where it is possible for the interference mechanism of stabilization to materialize. The possibility of making an adequate comparison of the numerical calculations with the results of a theoretical analysis of the interference stabilization phenomenon is due primarily to Ref. 9, where an analytical solution of the time-dependent Schrödinger equation describing the space–time evolution of the electronic wave function in a strong electromagnetic field and the dynamics of the ionization process, with the initial conditions for a Rydberg hydrogen atom was obtained on the basis of the WKB method.

In the present paper calculations of the ionization dynamics of Rydberg hydrogen atom are performed using an algorithm for direct numerical integration of the three-dimensional time-dependent Schrödinger equation with the linearly polarized field of an electromagnetic wave, for different values of the intensity and frequency of the laser radiation as well as different quantum numbers characterizing the initial state of the system. The probabilities for filling different states of the discrete spectrum and for ionization toward the end of the laser action are calculated. The ionization suppression found in the calculations is explained on the basis of the interference mechanism of stabilization, proposed in Ref. 4 and 7–9.

2. MODEL OF A THREE-DIMENSIONAL HYDROGEN-LIKE SYSTEM

The Hamiltonian of the quantum system was chosen in the form

$$H_0=-\frac{1}{2}\nabla^2-\frac{1}{\sqrt{\alpha^2+r^2}}, \tag{2}$$

where ∇^2 is the Laplacian, r is the magnitude of the radius vector, and α is a parameter smoothing the potential.

The stationary states of a system with the Hamiltonian (2) in a spherical coordinate system (r,θ,φ) are determined by the expression

$$\Phi_{nlm}(r,\theta,\varphi)=R_{nl}Y_{lm}(\theta,\varphi),$$

where $R_{nl}(r)$ are the radial wave functions, $Y_{lm}(\theta,\varphi)$ are spherical harmonics, and (n,l,m) is a set of quantum numbers describing the state of the electron.

The radial wave function is a solution of the eigenvalue problem

$$-\frac{1}{2}\frac{1}{r}\frac{\partial^2}{\partial r^2}rR_{nl}+V_{\text{eff}}(r)R_{nl}=E_{nl}R_{nl}, \tag{3}$$

where

$$V_{\text{eff}}=\frac{l(l+1)}{2r^2}-\frac{1}{\sqrt{\alpha^2+r^2}}$$

is the effective potential and E_{nl} are the energy levels of the system.

In the case $\alpha=0$ the equation (3) possesses an analytical solution, where $E_{nl}=-1/2n^2$; for $\alpha\neq 0$ the problem (3) must be solved numerically and the degeneracy with respect to the orbital angular momentum is lifted.

In our calculations we assumed $\alpha=1/2$ (in atomic units). This choice of α, on the one hand, preserves the structure of the spectrum of a real hydrogen atom and, on the other, greatly simplifies the numerical calculations near the point $r=0$. The results of the numerical calculations of the wave functions of several stationary states are presented in Fig. 1. Table I gives the values of the energy levels of a model system as well as of a real hydrogen atom. As one can see, the largest difference is observed for the $1s$ ground state and the lower excited states $2s$, $2p$. For states with $n\geqslant 3$ the energy levels are degenerate with respect to the orbital angular momentum and their energies equal the energies of the states of a real hydrogen atom to within $\leqslant 2\%$.

3. INTERACTION WITH THE FIELD OF AN ELECTROMAGNETIC WAVE

The interaction of a quantum system with the field of an electromagnetic wave is described in the dipole approximation by the operator

$$W=-\mathbf{d}\cdot\boldsymbol{\varepsilon}, \tag{4}$$

where $\boldsymbol{\varepsilon}(t)$ is the electric field vector of the wave and $\mathbf{d}$ is the dipole moment operator.

We shall assume that the electromagnetic wave is linearly polarized, and we orient the z axis along the vector $\boldsymbol{\varepsilon}$. Then

$$W=-z\varepsilon(t)=-r\varepsilon(t)\cos\theta. \tag{5}$$

Let the initial state of the atom be characterized by a zero projection of the orbital angular momentum on the z axis $(m=0)$. Since a linearly polarized field of an electromagnetic wave will give rise only to transitions without a change in the magnetic quantum number m $(\Delta m=0)$, the electron wave function is independent of the coordinate φ for any moment in time, and the evolution of the state of the atom in the field of the electromagnetic wave will be described by the two-dimensional time-dependent Schrödinger equation

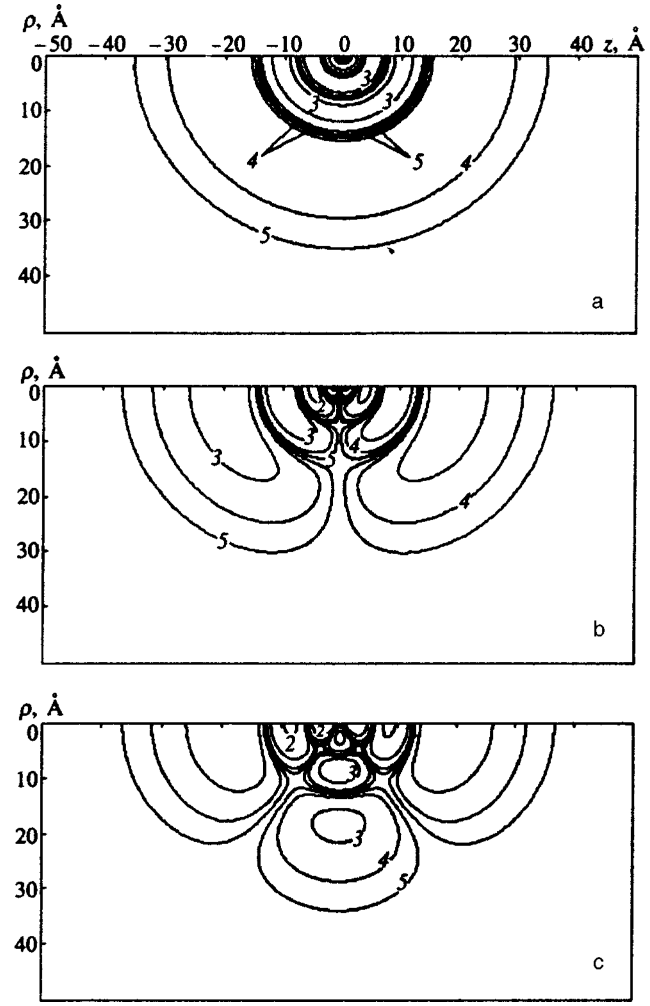


FIG. 1. Probability density $|\Phi_{nl}(\rho,z)|^2$ (where $z=r\cos\theta$, $\rho=r\sin\theta$) for the stationary states of a model hydrogen-like atom in the $5s$ (a), $5p$ (b), and $5d$ (c) states. Contour lines: *1*—10^{-2}, *2*—10^{-3}, *3*—10^{-4}, *4*—10^{-5}, *5*—10^{-6}.

$$i\frac{\partial\Psi}{\partial t}=-\frac{1}{2}\left[\frac{1}{r}\frac{\partial^2}{\partial r^2}r\Psi+\frac{1}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Psi}{\partial\theta}\right)\right]-\frac{1}{\sqrt{\alpha^2+r^2}}\Psi(r,\theta,t)-\varepsilon(t)\Psi(r,\theta,t)r\cos\theta. \tag{6}$$

We take as the initial condition a stationary state of the quantum system

$$\Psi(r,\theta,t=0)=\Phi_{nl}(r,\theta)=R_{nl}(r)Y_{l0}(\theta,\varphi)=R_{nl}(r)P_l(\cos\theta). \tag{7}$$

Here $P_l(\cos\theta)$ are the Legendre polynomials.

The calculations were performed for trapezoidal and square-wave pulses:

$$\varepsilon(t)=\begin{cases}\varepsilon_0\cos(\omega t)t/\tau_f, & 0<t\leq\tau_f,\\ \varepsilon_0\cos(\omega t), & \tau_f<t\leq\tau+\tau_f,\\ \varepsilon_0\cos(\omega t)\left[1-\dfrac{t-(\tau+\tau_f)}{\tau_f}\right], & \tau+\tau_f<t\leq\tau+2\tau_f,\end{cases} \tag{8}$$

$$\varepsilon(t)=\begin{cases}\varepsilon_0\cos(\omega t), & 0\leq t\leq\tau,\\ 0, & t>\tau.\end{cases} \tag{9}$$

Here τ_f and τ are the durations of the leading and trailing edges and of the ''shelf,'' and ε_0 is the amplitude of the electric field of the wave.

In the calculations the photon energy was varied in the range $\hbar\omega=1-5$ eV and the intensity in the range $P=3\cdot10^{12}-3\cdot10^{15}$ W/cm^2. As a rule, the times τ and τ_f were equal to $\tau_f=2T$ and $\tau=10T$, where $T=2\pi/\omega$ is the duration of the optical cycle.

4. COMPUTATIONAL PROCEDURE

The numerical integration of Eq. (6) was performed in a cylindrical coordinate system

$$z=r\cos\theta,\quad \rho=r\sin\theta$$

by the finite-element method using the procedure described in Ref. 21. For this, the wave function of the initial state, determined in spherical coordinates by means of Eq. (7), was discretized on a (ρ,z) grid.

The size of the computational region was chosen so as to be able to take into account at least 10 states with different values of the principal quantum number $n=1-n_{\max}$ ($n_{\max}=10$) and all orbital angular momentum states corresponding to a given value of n: $l=0-(n-1)$. Therefore $N=n_{\max}(n_{\max}+1)/2=55$ states in the discrete spectrum can be taken into account in the calculations. The control calcu-

TABLE I. Energy levels (eV) of a model hydrogen-like system.

l	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
0	−12.53	−3.319	−1.494	−0.843	−0.540	−0.376	−0.276	−0.211	−0.167	−0.134
1	-	−3.429	−1.530	−0.858	−0.548	−0.380	−0.279	−0.213	−0.168	−0.135
2	-	-	−1.525	−0.852	−0.545	−0.378	−0.278	−0.213	−0.168	−0.135
3	-	-	-	−0.848	−0.543	−0.377	−0.277	−0.212	−0.168	−0.135
4	-	-	-	-	−0.542	−0.377	−0.277	−0.212	−0.167	−0.135
5	-	-	-	-	-	−0.377	−0.277	−0.212	−0.167	−0.136
6	-	-	-	-	-	-	−0.277	−0.212	−0.167	−0.136
7	-	-	-	-	-	-	-	−0.212	−0.167	−0.136
8	-	-	-	-	-	-	-	-	−0.167	−0.136
9	-	-	-	-	-	-	-	-	-	−0.136
H	−13.61	−3.402	−1.512	−0.851	−0.544	−0.378	−0.277	−0.213	−0.168	−0.136

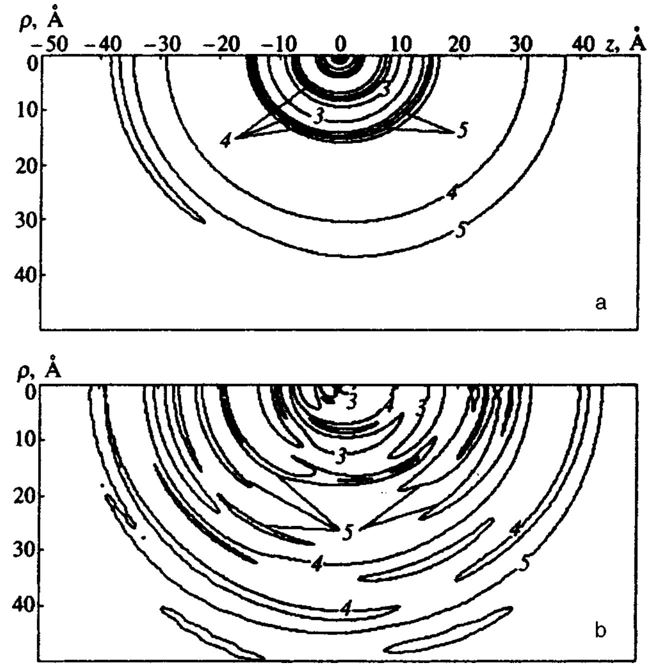


FIG. 2. Electron density distribution $|\Phi_{nl}(\rho,z)|^2$ by the end of a laser pulse with $\hbar\omega=5$ eV, $\tau_f=2T$, and $\tau=10T$ for initial state 5s and radiation intensity $P=10^{13}$ W/cm^2 (a) and $3\cdot10^{14}$ W/cm^2 (b). The contour lines correspond to Fig. 1.

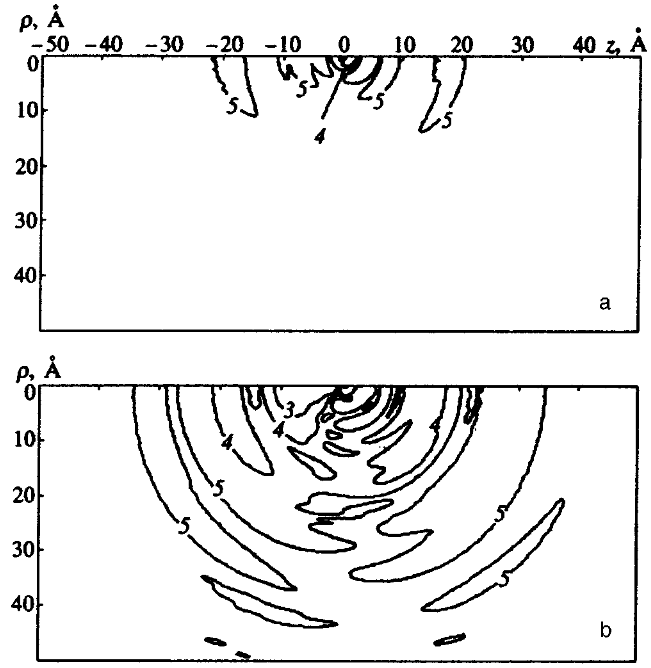


FIG. 3. Electron density distribution in the continuum for the parameters corresponding to Fig. 2.

lations showed that under the specified conditions the size of the spatial region of the calculation should be at least $r_{\max}=180$ Å For this reason, the Schrödinger equation was discretized on a rectangular (ρ,z) grid in the region

$$|z|\leqslant z_{\max}=180\ \text{Å},$$

$$0\leqslant\rho\leqslant\rho_{\max}=180\ \text{Å}.$$

The grid was constructed with a nonuniform step, so as to ensure the required computational accuracy near the point $r=0$ in the region of rapidly varying potential. The total number of points in the spatial grid was $\rho\times z=480\times960$. An imaginary component, ensuring ''swallowing'' of the wave function and absence of reflection from the boundaries, was introduced into the potential near the boundaries $\rho_{\max}$, $z_{\max}$ of the computational region.

The populations of different states of the discrete spectrum were found by projecting the computed wave function $\Psi(r,\theta,t)$ onto the basis of atomic states $\{R_{nl}(r)P_l(\cos\theta)\}$ which are not perturbed by the field of the electromagnetic wave:

$$W_{nl}(t)=|C_{nl}(t)|^2=\left|\int\Psi^*(r,\theta,t)R_{nl}(r)P_l(\cos\theta)r^2dr\,d\Omega\right|^2. \tag{10}$$

The ionization probability was calculated from the formula

$$W_i=1-\sum_{n=1}^{n_{\max}}\sum_{l=0}^{n-1}|C_{nl}(t)|^2. \tag{11}$$

The calculations were performed in the range of intensities for which the populations of states with $n>n_{\max}$ is negligibly small.

An alternative approach to calculating the ionization probability is to calculate the fraction of the electron probability density that is ''eaten away'' near the boundaries of the spatial region of the calculation at long times after laser action ceases. This approach was implemented in Refs. 16 and 22. However, it should be kept in mind that the choice of a quite small spatial region for the purpose of decreasing the computational time and economizing memory (in Ref. 16 the computational region did not exceed 75×20 Å^2) inevitably will result in a loss of information about the population of the excited states and will automatically overestimate of the ionization probability, especially in strong fields.

5. MODELING RESULTS

The bulk of the calculations were performed for a trapezoidal laser pulse with the parameters $\tau_f=2T$ and $\tau=10T$ for $\hbar\omega=5$ eV. In this case the region of strong fields $\varepsilon/\omega^{5/3}>1$ corresponds to radiation intensity $P>2.5\times10^{-3}$ a.u. ($\sim10^{14}$ W/cm^2). Stationary states with the quantum numbers $n=1\cdots7$, $l=0\cdots(n-1)$ were chosen as the initial states. For all states with $n\geqslant2$ ionization was a single-photon process.

Figure 2 shows the typical electron density distribution $|\Psi(\rho,z)|^2$ by the end of the laser pulse for different values of the radiation intensity and for the initial state 5s. Figure 3 shows the electron density distribution $\rho_c=|\Psi_c(\rho,z)|^2$ in the continuum for the same conditions as for the data in Fig. 2. These distributions were calculated as follows:

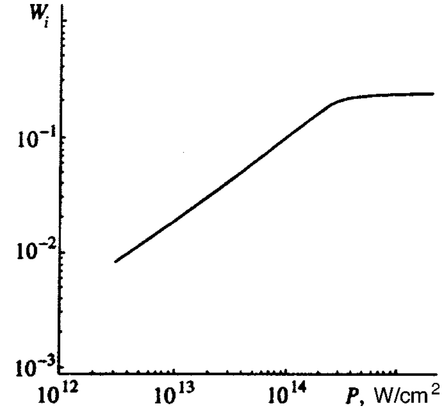


FIG. 4. Ionization probability for a hydrogen atom in the $5s$ state over the pulse duration versus the radiation intensity for $\hbar\omega=5$ eV, $\tau_f=2T$, and $\tau=10T$.

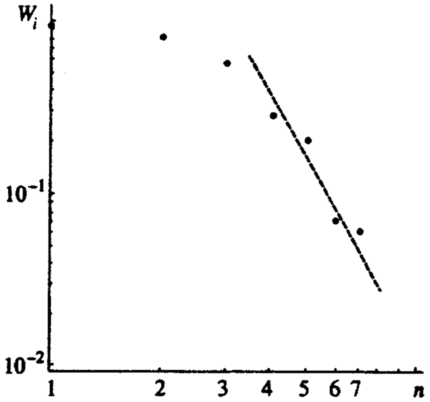


FIG. 5. Ionization probability of a hydrogen atom in the s state over the pulse duration versus the initial value of the principal quantum number for $P=3\times10^{14}$ W/cm^2, $\hbar\omega=5$ eV, $\tau_f=2T$, and $\tau=10T$. Dashed curve—$W_i\propto n^{-3}$.

$$\rho_c(\rho,z,t)=|\Psi_c(\rho,z,t)|^2=\left|\Psi(\rho,z,t)-\sum_{n,l}C_{nl}(t)\Phi_{nl}(\rho,z)\exp(-iE_{nl}t)\right|^2, \quad (12)$$

where $|C_{nl}|^2$ is the probability, determined using the relation (10), of observing the system in the state nl of the discrete spectrum. We note that as the radiation intensity increases, the degree of spherical symmetry of the wave packet in the continuum increases, possibly as a result of an increase in the role of multiphoton processes and filling of continuum states with different values of l.

Figure 4 shows how the ionization probability W_i of a hydrogen atom from the initial $5s$ state depends on the radiation intensity in the range $P=3\times10^{12}-3\times10^{15}$ W/cm^2 for $\hbar\omega=5$ eV, $\tau_f=2T$, and $\tau=10T$. As one can see, in the region of weak fields W_i increases linearly with the intensity of the wave, and above a critical value corresponding to the condition (1) it reaches a plateau, i.e., the system stabilizes. A similar dependence $W_i(P)$ is predicted in the latest models in the theory of interference stabilization.[8,9]

To determine the nature of the stabilization mechanism we investigated the ionization probability as a function of the quantum numbers n and l characterizing the initial state of the atom. The dependence of the ionization probability over a pulse on the principal quantum number of the initial s state was obtained for radiation intensity $P=3\times10^{14}$ W/cm^2 ($\varepsilon/\omega^{5/3}=1.57$). The data from the corresponding calculations are presented in Fig. 5. As one can see, in the region $n\geqslant4$ the function $W_i(n)$ is approximated well by the expression

$$W_i(n)\propto n^{-3}. \quad (13)$$

Such a $W_i(n)$ dependence agrees with the theory of interference stabilization of Rydberg atoms (see, for example, Ref. 9). The computational results also permit asserting that the function (13) obtained for Rydberg states $n\gg1$ is actually valid even for states with relatively small values of the principal quantum number. However, it should be noted that a similar function $W_i(n)$ is also obtained for weak fields on the basis of Fermi's golden rule, since in this case[6]

$$\dot{W}_i\propto|V_{nE}|^2\propto1/n^3. \quad (14)$$

Here $\dot{W}_i$ is the ionization probability per unit time.

The ionization probability was also studied as a function of the orbital angular momentum quantum number of the initial state of an electron for $n=5$ (see Fig. 6). It was found that the ionization stability of the system increases with l. The reason for this is that the centrifugal potential barrier, which ‘‘pries’’ the electron off the nucleus for larger values of r, increases and hence the probability of absorption of photons from the field decreases. The increase in the ionization stability of states with a large orbital angular momentum was discussed earlier in Ref. 23. We also note that in the quasiclassical theory[9] the centrifugal potential is assumed to be weak, compared with the interaction of the electron with the nucleus, for values of l satisfying the condition

$$l<\omega^{-1/3}. \quad (15)$$

For $\hbar\omega=5$ eV ($\omega\cong0.18$ a.u.) the condition (15) holds for the s and p states. As one can see from the data presented in

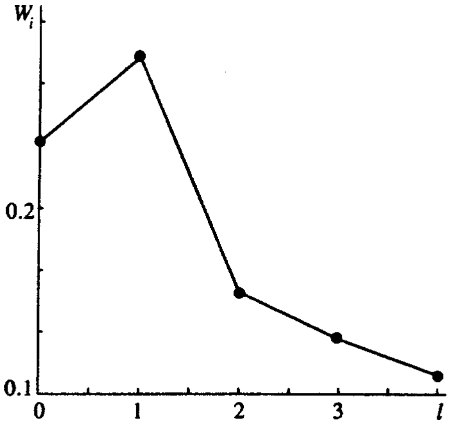


FIG. 6. Ionization probability of a hydrogen atom over the pulse duration versus the orbital angular momentum quantum number for an initial state with $n=5$. The parameters of the laser pulse are the same as for Fig. 5.

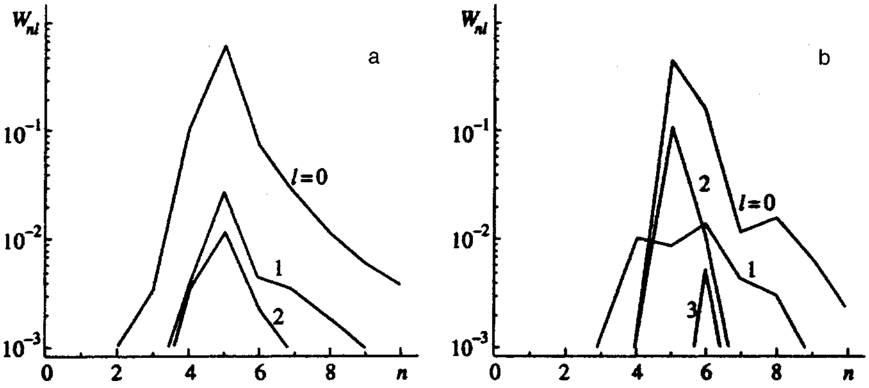


FIG. 7. Probability distribution for filling of different states in the discrete spectrum by the end of the laser pulse. The duration of the "shelf" was $\tau=5T$ (a) and $20T$ (b). Radiation intensity was $P=3\cdot 10^{14}$ W/cm^2, $\hbar\omega=5$ eV.

Fig. 6 the ionization probabilities of the $5s$ and $5p$ states are indeed close to one another, and the effect of the centrifugal potential is to decrease the ionization probability of states with $l\geqslant 2$.

The probability density for filling of different states in the discrete spectrum by the end of the laser action is of special interest. The results of the corresponding calculations for the initial $5s$ state and laser radiation parameters $P=3\cdot 10^{14}$ W/cm^2 and $\hbar\omega=5$ eV are presented in Fig. 7. The data obtained show that the s states—mainly the $4s$ and $6s$ states which are coupled with the initial $5s$ state by Λ-type two-photon transitions—are predominantly filled. Moreover, there exists a nonzero probability of nonresonant single-photon filling of the $5p$ state. For short laser pulse durations ($\tau=5T$) it equals 2.7% by the end of the laser pulse, but it is higher than the probability of a resonant two-photon Λ-transition $5s\rightarrow 5d$ (see Fig. 7a). As the duration of the laser pulse increases to $\tau=20T$ (see Fig. 7b) the probability of the transition $5s\rightarrow 5d$ increases substantially, while the probabilities of nonresonant $5s\rightarrow np$ transitions remain at the 1–2% level.

These calculations showed that for the initial $5s$ state under conditions when the resonant two–photon Λ-transitions predominate over nonresonant single-photon transitions the transitions $5s\rightarrow 4s$, $6s$, $5d$ are strongest in the entire range of intensities investigated.

However, the role of states with a nonzero orbital angular momentum in the dynamics of the ionization of a Rydberg atom and the relation of the filling of these states with the stabilization phenomenon and the form of the function $W_i(P)$ are still not completely understood. For example, it remains unclear whether or not the absence of the "valley of death" predicted in Ref. 4 is due to the fact that different orbital angular momentum states are taken into account[8] or to the failure of a number of approximations, such as the pole approximation and the rotating-wave approximation,[9] which were used in Refs. 4 and 7. These questions all require additional study.

It is also of interest to investigate the frequency dependence of the ionization probability of an excited hydrogen atom. It should be expected that under the conditions for which the method developed in Ref. 9 is applicable the ionization probability over a laser pulse of fixed duration will remain unchanged provided that the frequency and intensity of the radiation are chosen so that

$$V=\varepsilon/\omega^{5/3}=\text{const.} \tag{16}$$

Calculations for different values of the radiation frequency in the range $\hbar\omega=1-5$ eV (the corresponding radiation intensities varied in the range $1.41\times 10^{12}-3\times 10^{14}$ W/cm^2) were performed for a 12.4 fs (which corresponds to 15 optical cycles with $\hbar\omega=5$ eV) square pulse. For $\hbar\omega=1$ eV the pulse duration was equal to only three optical cycles. The computational results for $W_i(\omega,\varepsilon/\omega^{5/3}=\text{const})$ are presented in Fig. 8 and show that the similarity criterion with respect to the parameter (16) holds to a quite high degree of accuracy. This fact likewise supports the interference mechanism of stabilization in our range of parameter values.

6. CONCLUSIONS

The calculations of the ionization dynamics of an excited hydrogen-like atom performed in this work showed stabilization for intensities in the range $P\geqslant 2\times 10^{14}$ W/cm^2 with $\hbar\omega=5$ eV. It was proved on the basis of an investigation of the dependence of the ionization probability on the quantum numbers of the initial state of the atom and on the parameter $\varepsilon/\omega^{5/3}$ as well as on the basis of the data obtained on the filling of different states in the discrete spectrum by the end of the pulse that the observed stabilization is of an interference nature. It should be noted that in the case studied a Rydberg hydrogen atom can be understood to be an atom in an excited state with a relatively small principal quantum number $n\geqslant 4$.

We believe that it would be extremely important and interesting to investigate the possibility of observing interference stabilization in a three-dimensional quantum system with a short-range potential, when the number of states in the

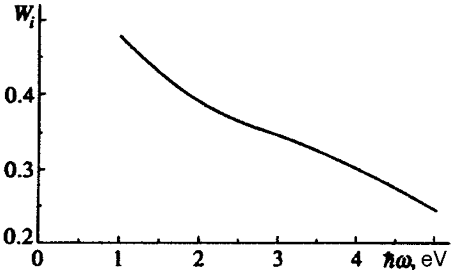


FIG. 8. Ionization probability versus radiation frequency in a 12.4 fs square pulse with a fixed value of the parameter $\varepsilon/\omega^{5/3}=1.57$.

discrete spectrum is finite, and the degeneracy of the states with respect to the orbital angular momentum is lifted.

The question of the connection between interference stabilization of an atom and stabilization in the Kramers–Henneberger regime also remains unsolved. In Ref. 24 it was shown that stabilization in the Kramers–Henneberger regime can also arise in a situation when there is one state in the discrete spectrum. This circumstance attests convincingly to the fact that the physical mechanisms of both phenomena are different and one phenomenon does not reduce to the other. At the same time, the question arises of whether both stabilization mechanisms can exist simultaneously in superstrong fields under conditions when a double-well potential describing a ‘‘dressed’’ atom is formed, and also the question of the competition between these mechanisms.

We thank M. V. Fedorov for a discussion of the formulation of the problem as well as a discussion of the results obtained.

This work was supported by the Russian Fund for Fundamental Research (Project No. 96-02-19286).

*)E-mail: popov@mics.msu.su

---

[1] M. V. Fedorov, *Interaction of Intense Laser Light with Free Electrons* [in Russian], Nauka, (Harwood Academic, New York, 1991).
[2] N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms*, Springer, Heidelberg, 1993.
[3] N. B. Delone and V. P. Kraĭnov, Usp. Fiz. Nauk **165**, 1295 (1995).
[4] M. V. Fedorov and A. M. Movsesian, J. Phys. B **21**, L155 (1988).
[5] M. Gavrila and J. Kaminski, Phys. Rev. Lett. **52**, 613 (1984).
[6] S. P. Goreslavskiĭ, N. B. Delone, and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **82**, 1789 (1982) [Sov. Phys. JETP **55**, 1032 (1982)].
[7] M. V. Fedorov, J. Phys. B **27**, 4145 (1994).
[8] M. V. Fedorov, M.-M. Tehranchi, and S. M. Fedorov, J. Phys. B **29**, 2907 (1996).
[9] O. V. Tikhonova and M. V. Fedorov, Laser Phys. **7**, 574 (1997).
[10] Q. Su, J. Eberly, and J. Javanainen, Phys. Rev. Lett. **64**, 861 (1990).
[11] R. Grobe and M. V. Fedorov, Laser Phys. **3**, 265 (1993).
[12] W. Su, A. Sanpera, and L. Roso-Franco, Int. J. Modern Phys. **8**, 1655 (1994).
[13] A. M. Popov, O. V. Tikhonov, and E. A. Volkova, Laser Phys. **5**, 1184 (1995).
[14] E. A. Volkova, A. M. Popov, and O. V. Tikhonova, Zh. Éksp. Teor. Fiz. **109**, 1586 (1996) [JETP **82**, 853 (1996)].
[15] A. M. Popov, O. V. Tikhonova, and E. A. Volkova, Laser Phys. **7**, 329 (1997).
[16] K. C. Kulander, K. J. Shafer, and J. L. Krause, Phys. Rev. Lett. **20**, 2601 (1991).
[17] M. Gajda, B. Piraux, and K. Rzazewski, Phys. Rev. A **50**, 2528 (1994).
[18] R. R. Jones and P. H. Bucksbaum, Phys. Rev. Lett. **67**, 3215 (1991).
[19] R. R. Jones, D. Schumacher, and P. Bucksbaum, Phys. Rev. A **47**, R49 (1993).
[20] J. Hoogenraad, R. Vrijen, and L. Noordam, Phys. Rev. A **50**, 4133 (1994).
[21] E. A. Volkova, A. M. Popov, and O. V. Tikhonova, Zh. Éksp. Teor. Fiz. **108**, 436 (1995) [JETP **81**, 235 (1995)].
[22] K. C. Kulander, Phys. Rev. A **35**, 445 (1987).
[23] H. G. Muller and H. B. van Linden van den Heuvell, Laser Phys. **3**, 694 (1993).
[24] E. A. Volkova, A. M. Popov, O. V. Tikhonova, and O. V. Smirnova, Zh. Éksp. Teor. Fiz. **111**, 1194 (1997) [JETP **84**, 658 (1997)].

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Integrals of the motion and exact solutions of the problem of two dispersing delta-wells

V. I. Man'ko*[)]

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

A. S. Chikhachev

Salyut Design Bureau, M. V. Khrunichev State Space Scientific-Production Center, 101000 Moscow, Russia

(Submitted 5 September 1997)
Zh. Éksp. Teor. Fiz. **113**, 606–614 (February 1998)

An exact solution is analyzed for the analogs of bound and scattering states in a nonstationary quantum mechanical system whose potential has the form of two dispersing delta-wells. For the delta-potentials explicit (in the form of operator kernels) expressions are found for the integrals of the motion that depend on time and transform to the known integrals of the motion for a free quantum particle as the interaction force with the potential approaches zero. © *1998 American Institute of Physics.* [S1063-7761(98)01702-8]

1. Exact solutions of the Schrödinger equation for time-dependent quantum mechanical Hamiltonians exist only for a few potentials. Thus, a variable frequency oscillator was examined, and its wave functions (both Gaussian packets and discrete analogs of bound states) found explicitly by Husimi.[1] A propagator for this problem in a Gaussian form was also obtained there. The problem for a constant frequency oscillator acted on by a variable external force was solved independently by Schwinger[2] and Feynman.[3] An integral of the motion that was quadratic in the coordinates and momentum and has the same form as an integral of motion found by Ermakov[4] for the classical parametric oscillator was constructed by Lewis and Riesenfeld.[5,6] Linear integrals of the motion for the quantum parametric oscillator have been constructed for the one-dimensional[7,8] and multidimensional[9] cases. Other nonstationary problems for which an exact solution has been obtained include cyclotron motion in a variable field,[10] as well as problems with specially selected time dependences.[11]

The delta-potential is a good model for short- range forces. A propagator for a stationary delta-well has been obtained as a function of the error function.[12,13] The propagator for a single nonstationary delta-well has also been examined.[14] There is great interest in a model for to dispersing delta-wells. Exact solutions were first obtained by Zhdanov and Chikhachev[15] who constructed a symmetric solution that transforms to the known bound state for a zero dispersion velocity and analogs of scattering states were obtained later.[16] For nonquadratic potentials such as the delta-wells, however, the integrals of the motion were unknown. The purpose of this paper is to study in detail the properties of the solutions for two dispersing delta-wells, including the antisymmetric solution. An explicit, time dependent expression for the integral of the motion is also found for a single delta-well and the limiting transition to the case of free motion is studied.

2. Let us first consider the state of a particle with a single attractive delta-center at the point $x=0$. The equation for the Green function G in this case has the form

$$\left[i\frac{\partial}{\partial t}+\frac{1}{2}\frac{\partial^2}{\partial x^2}+\alpha\delta(x)\right]G(x,x',t,t') = i\delta(x-x')\delta(t-t'). \tag{1}$$

Using a representation of G in the form of the sum

$$\sum_\lambda \psi_\lambda^*(x',t')\psi_\lambda(x,t),$$

where the ψ_λ are the eigenfunctions of the operator

$$i\frac{\partial}{\partial t}+\frac{1}{2}\frac{\partial^2}{\partial x^2}+\alpha\delta(x),$$

in a fashion similar to that used in Ref. 13, we can obtain

$$\begin{aligned} G={}&\frac{1}{\sqrt{2\pi i(t-t')}}\exp\left[i\frac{(x-x')^2}{2(t-t')}\right]\sigma(t-t')+\alpha \\ &\times\exp\left[-\alpha(|x|+|x'|)+i\frac{\alpha^2(t-t')}{2}\right] \\ &-\frac{\alpha}{\pi}\int_0^\infty\frac{dk}{k^2+\alpha^2}\{k\sin[k(|x|+|x'|)] \\ &+\alpha\cos[k(|x|+|x'|)]\}=G_0+\alpha \\ &\times\exp\left[-\alpha(|x|+|x'|)+\frac{i\alpha^2(t-t')}{2}\right] \\ &-\frac{\alpha i}{2\pi}\int_{-\infty}^{\infty}\frac{dk}{k+i\alpha} \\ &\times\exp\left[-ik^2\frac{t-t'}{2}-ik(|x|+|x'|)\right], \end{aligned} \tag{2}$$

where

$$\sigma(\tau)=1,\quad \tau\geqslant 0,\quad \sigma=0\quad\text{for }\tau<0.$$

The first term in the expression for G is the propagator for a free particle, the second terms corresponds to the single bound state of the particle, and the integral term is the sum over all the continuum states.

The integral term can be expressed in terms of the Moshinsky function:[17]

$$M(x,k,t)=\frac{i}{2\pi}\int_{-\infty}^{\infty}\frac{dk'}{k-k'}\exp\left[i\left(k'x-\frac{k'^2t}{2}\right)\right].$$

Transforming this integral using the results of Ref. 17 makes it possible to write the Green function in the form

$$G(x,x',t)=G_0(x,x',t)+\alpha M_1(\theta,\alpha,t),\tag{3}$$

where

$$G_0=\frac{1}{\sqrt{2\pi it}}\exp\left[i\,\frac{(x-x')^2}{2t}\right]\sigma(t),$$

$$M_1=\frac{1}{2}\exp\left(-\alpha\theta+\frac{i\alpha^2t}{2}\right)\operatorname{erf}\left(\exp\left(\frac{3\pi i}{4}\right)\frac{\theta-i\alpha t}{\sqrt{2t}}\right),$$

$$\operatorname{erf}(z)=\frac{2}{\sqrt{\pi}}\int_0^z\exp(-x^2)dx,\quad \theta=|x|+|x'|.$$

Equation (3) for the Green function differs from that derived in Ref. 17, since the case $\alpha<0$, i.e., without a bound state, was discussed there.

The explicit expressions (2) and (3) for the propagator also make it possible to find the integrals of motion for our problem explicitly. In fact, it has been shown[18] that, if the evolution operator $\hat{U}(t)$ of the system is given, i.e.,

$$\hat{U}(t,t')\Psi(x,t')=\Psi(x,t),\tag{4}$$

where $\Psi(x,t')$ is the initial value of the wave function at time t', then the operator

$$\hat{I}(t,t')=\hat{U}(t,t')I(t')\hat{U}^{-1}(t,t')\tag{5}$$

is an integral of the motion. This means that $d\hat{I}/dt\equiv 0$. The Green function, by definition, is a matrix element of the evolution operator,

$$G(x,x',t,t')=\langle x|\hat{U}(t,t')|x'\rangle$$

in the coordinate representation, i.e., the kernel of the evolution operator.

We are examining a system with a hermitian Hamiltonian, so that the evolution operator is unitary. Equation (5) implies that if the kernel of the operator $\hat{I}(t')$ in the coordinate representation, i.e.,

$$I(x,x',t')=\langle x|\hat{I}(t')|x'\rangle,$$

is known, then for the kernel of the integral of motion $\hat{I}(t,t')$ in this representation we have the expression

$$J(x,x',t,t')=\int G(x,z,t,t')I(t')G^*(z,x',t,t')dz.$$

For concreteness we set $t'=0$ and consider separately the cases of a coordinate and a momentum as the operator $\hat{I}|_{t'=0}$.

For the coordinate operator, we obtain the equation

$$\begin{aligned}(\hat{x}_0(t))_{xx'}&=\int_{-\infty}^{\infty}z\,dz\,G(x,z,t)G^*(z,x',t)\\&=\frac{t}{2\pi}\,\delta'(x-x')\exp\left(i\,\frac{x^2-x'^2}{2t}\right)\\&\quad-i\alpha\int_0^{\infty}z\,dz\,\sin\frac{xz}{t}\exp\left(i\,\frac{z^2}{2t}-\alpha z\right)\\&\quad\times\operatorname{erf}\left(\exp\left(\frac{3\pi i}{4}\right)\frac{z+|x'|-i\alpha t}{\sqrt{2t}}\right)\\&\quad\times\exp\left(i\,\frac{x^2}{2t}-\alpha\left|x'\right|+\frac{i\alpha^2t}{2}\right)\\&\quad+i\alpha\int_0^{\infty}z\,dz\,\sin\frac{x'z}{t}\exp\left(-\frac{iz^2}{2t}-\alpha z\right)\\&\quad\times\operatorname{erf}\left(\exp\left(-\frac{3\pi i}{4}\right)\frac{z+|x|+i\alpha t}{\sqrt{2t}}\right)\\&\quad\times\exp\left(-\frac{ix^2}{2t}-\alpha\left|x\right|-\frac{i\alpha^2t}{2}\right).\end{aligned}\tag{6}$$

Similarly, for the momentum, we can obtain the formula

$$\begin{aligned}(\hat{p}_0(t))_{xx'}&=i\int_{-\infty}^{\infty}dz\,G(x,z,t)\,\frac{\partial}{\partial z}\,G^*(z,x',t)\\&=-\frac{x'}{2\pi}\,\delta(x-x')+\frac{i}{2\pi}\,\delta'(x-x')\\&\quad\times\exp\left(i\,\frac{x^2-x'^2}{2t}\right)\\&\quad+\alpha\left\{\frac{1}{\sqrt{-2\pi it}}\int_0^{\infty}dz\,M_1(|x|+z,\alpha,t)\right.\\&\quad\times\left(\frac{x'}{2t}\sin\frac{x'z}{2t}-\frac{z}{2t}\cos\frac{x'z}{2t}\right)\\&\quad\times\exp\left(\frac{ix'^2}{2t}+\frac{iz^2}{2t}\right)-\frac{1}{\sqrt{2\pi it}}\int_0^{\infty}dz\\&\quad\times M_1^*(|x'|+z,\alpha,t)\left(\frac{x}{2t}\sin\frac{xz}{2t}-\frac{z}{2t}\cos\frac{xz}{2t}\right)\\&\quad\left.\times\exp\left(-\frac{ix^2}{2t}-\frac{iz^2}{2t}\right)\right\}.\end{aligned}\tag{7}$$

In Eqs. (6) and (7) the terms proportional to α are responsible for the difference between the analogous expressions for free particle motion. We note also that arbitrary functions of $\hat{p}_0$ and $\hat{x}_0$ are also integrals of motion. On the other hand, any other integrals of the motion, such as $E=p^2/2+\alpha\delta(x)$, can be expressed in terms of $\hat{p}_0$ and $\hat{x}_0$.

3. In the case of dispersing delta-wells, the system of eigensolutions consists of two ''bound'' states, or states described by an exponentially decaying functions, and states with an oscillating asymptote.

As Dappen[19] has done, we shall seek decaying solutions that are symmetric and antisymmetric in x to the equation

$$\left\{ i\frac{\partial}{\partial t}+\frac{1}{2}\frac{\partial^2}{\partial x^2}+\alpha[\delta(x-vt)+\delta(x+vt)]\right\}\Psi(x,t)=0. \tag{8}$$

For the symmetric solution we set

$$\begin{aligned}\Psi^{(+)}=\exp\left(i\frac{v^2t}{3}\right)\sum_{s=0}^{\infty}C_s^{(+)}\{&\exp[iv(x-vt)\\ &-(\alpha+2ivs)|x-vt|]+\exp[-iv(x+vt)\\ &-(\alpha+2ivs)|x+vt|]\}\exp\left[\frac{it}{2}(\alpha+2ivs)^2\right].\end{aligned} \tag{9}$$

If Eq. (9) satisfies Eq. (8), then the $C_s^{(+)}$ can be represented in the form

$$C_s^{(+)}=\frac{1}{s!}\left(\frac{\alpha}{2iv}\right)^s\sqrt{\frac{\alpha}{2}}\exp\left(\frac{i\alpha}{2v}\right).$$

Here, for $v\to 0$, the solution (8) transforms into a solution of the equation

$$\left[i\frac{\partial}{\partial t}+\frac{1}{2}\frac{\partial^2}{\partial x^2}+2\alpha\delta(x)\right]\Psi(x,t)=0,$$

corresponding to a bound state. (See Ref. 15.)

For the solution antisymmetric in x,

$$\begin{aligned}\Psi^{(-)}=\exp\left(i\frac{v^2t}{2}\right)\sum_{s}C_s^{(-)}\{&\exp[iv(x-vt)\\ &-(\alpha+2ivs)|x-vt|]-\exp[-iv(x+vt)\\ &-(\alpha+2ivs)|x+vt|]\}\exp\left[\frac{it}{2}(\alpha+2ivs)^2\right],\end{aligned}$$

$$C_s^{(-)}=\frac{1}{s!}\left(-\frac{\alpha}{2iv}\right)^s\sqrt{\frac{\alpha}{2}}\exp\left(-\frac{i\alpha}{2v}\right). \tag{10}$$

The expressions for $C_s^{(+)}$ and $C_s^{(-)}$ differ from the corresponding expressions in Ref. 19 only by the factors $\exp(i\alpha/2v)$ and $\exp(-i\alpha/2v)$.

These factors (which do not affect the probability of charge exchange processes, have been calculated elsewhere[19]) are introduced in order to ensure a smooth transition to the case $v\to 0$.

At $t=0$ the series for $\Psi^{(\pm)}$ are easily summed as series for the exponentials:

$$\begin{aligned}\Psi^{(+)}|_{t=0}&=\sum_s\frac{1}{s!}\left(\frac{\alpha}{2iv}\right)^s\sqrt{\frac{\alpha}{2}}\exp\left(\frac{i\alpha}{2v}\right)\\ &\quad\times[\exp(ivx-\alpha|x|-2ivs|x|)\\ &\quad+\exp(-ivx-\alpha|x|-2ivs|x|)]\\ &=\sqrt{\frac{\alpha}{2}}\exp\left(\frac{i\alpha}{2v}\right)\\ &\quad\times\exp(-\alpha|x|)2\cos(vx)\\ &\quad\times\exp\left[\frac{\alpha}{2iv}\exp(-2iv|x|)\right],\end{aligned} \tag{11}$$

and

$$\begin{aligned}\Psi^{(-)}|_{t=0}=2i\sqrt{\frac{\alpha}{2}}\exp\left(-\frac{i\alpha}{2v}\right)&\exp(-\alpha|x|)\sin(vx)\\ &\times\exp\left[-\frac{\alpha}{2iv}\exp(-2iv|x|)\right].\end{aligned} \tag{12}$$

Equations (11) and (12) show that the important singularity at $v\to 0$ vanishes after introducing the factors $\exp(\pm i\alpha/2v)$.

For the continuum, which is characterized by oscillatory asymptotic behavior and corresponds to free motion with momentum k, as in Ref. 13, we write the ψ-function in the form of a sum,

$$\psi_k=\tilde{\psi}_k+\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{ik^2t}{2}+ikx\right), \tag{13}$$

and in the following we calculate separately the parts that are symmetric and antisymmetric in x:

$$\begin{aligned}\psi_k^{(+)}&=\tilde{\psi}_k^{(+)}+\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{ik^2t}{2}\right)\cos(kx),\\ \psi_k^{(-)}&=\tilde{\psi}_k^{(-)}+\frac{i}{\sqrt{2\pi}}\exp\left(-\frac{ik^2t}{2}\right)\sin(kx).\end{aligned} \tag{14}$$

As in Ref. 15, the functions $\tilde{\psi}_k^{(\pm)}$ can be sought in the form of the sums

$$\begin{aligned}\tilde{\psi}_k^{(\pm)}=\exp\left(\frac{iv^2t}{2}\right)\sum_s C_s^{(\pm)}[&\exp(ivx)\varphi_s(z_-t)\\ &\pm\exp(-ivx)\varphi_s(z_+t)],\end{aligned} \tag{15}$$

where

$$\varphi_s(z,t)=\exp\left(-a_sz+\frac{ia_s^2t}{2}\right),\quad z_-=|x-vt|,$$

$$z_+=|x+vt|.$$

For the coefficients $C_s^{(\pm)}$ we obtain the system

$$\begin{aligned}\sum_s C_s^{(\pm)}&\left\{(\alpha-a_s)\exp\left(i\frac{a_s^2t}{2}\right)\pm\alpha\right.\\ &\left.\times\exp\left[\frac{it}{2}(a_s+2iv)^2\right]\right\}\\ &=\frac{\alpha}{\sqrt{2\pi}}\left\{\exp\left[-\frac{it}{2}(k-v)^2\right]\right.\\ &\left.\pm\exp\left[-\frac{it}{2}(k+v)^2\right]\right\}.\end{aligned}$$

This system is equivalent to the following:

$$\sum_s \left\{ C_s^{(\pm)}(\alpha-a_s)\exp\left(i\,\frac{a_s^2 t}{2}\right) \pm \alpha C_{s-1}^{(\pm)} \times\exp\left[\frac{it}{2}\,(a_{s-1}+2iv)^2\right]\right\} = -\frac{\alpha}{2\sqrt{2\pi}}\left\{\exp\left[-\frac{it}{2}\,(k-v)^2\right] \pm \exp\left[-\frac{it}{2}\,(k+v)^2\right]\right\}. \tag{16}$$

To determine the coefficients $C_s^{(\pm)}$ we set $a_s=i[-|k|+v(2s-1)]$, with $a_0=-i(|k|+v)$, $a_1=-i(|k|-v)$, and $C_{-1}^{(\pm)}=0$. Then,

$$C_0^{(\pm)}=-\frac{\alpha}{2\sqrt{2\pi}}\,\frac{1}{\alpha-a_0},\quad C_1^{(+)}=-\frac{\alpha a_0}{2\sqrt{2\pi}(\alpha+a_0)},$$

and for $s>1$

$$C_s^{(+)}(\alpha-a_s)+\alpha C_{s-1}^{(+)}=0,$$

$$C_s^{(+)}|_{s\geqslant 1}=-\frac{a_0}{2\sqrt{2\pi}}\,\frac{1}{2iv}\left(\frac{\alpha}{2iv}\right)^s\frac{\Gamma(-\xi)}{\Gamma(s-\xi+1)}, \tag{17}$$

where

$$\xi=\frac{\alpha}{2iv}+\frac{k}{2iv}+\frac{1}{2iv}=\frac{\alpha-a_0}{2iv}.$$

The coefficients $C_s^{(-)}$ of the antisymmetric solution are determined by the equations

$$C_s^{(-)}=(-1)^s C_s^{(+)}, \tag{18}$$

and in the limit $v\to 0$ the solution for a single δ-center, characterized by the constant 2α, can be obtained. For $\nabla=0$,

$$C_s^{(+)}|_{s\geqslant 1}=-\frac{a_0}{2\sqrt{2\pi}}\,\frac{(-\alpha)^s}{(\alpha+i|k|)^{s+1}},$$

$$C_0^{(\pm)}=-\frac{\alpha}{2\sqrt{2\pi}}\,\frac{1}{\alpha+i|k|},$$

and the series $\psi_k^{(+)}$ is easily summed:

$$\tilde{\psi}_k^{(+)}=\frac{2i\alpha\cos(|k|x)}{i|k|+2\alpha}\,\exp\left(-\frac{ik^2t}{2}\right).$$

For $k<0$ the function $\psi_k^{(+)}$ does not change ($\tilde{\psi}_k^{(+)}=\tilde{\psi}_{-k}^{(+)}$), while $\tilde{\psi}_k^{(-)}$ changes sign ($\tilde{\psi}_k^{(-)}=-\tilde{\psi}_{-k}^{(-)}$).

The system of eigenfunctions determined by Eqs. (9), (10), and (14) probably forms a complete system.

4. We now study the generalization of Eq. (8)–dispersing delta-centers characterized by a different depth of the single bound state:

$$\left[i\,\frac{\partial}{\partial t}+\frac{1}{2}\,\frac{\partial^2}{\partial x^2}+\alpha\delta(x-vt)+\beta\delta(x+vt)\right]\Psi=0. \tag{19}$$

Here we consider a "bound" state of a nonstationary system corresponding to an asymptote that falls off exponentially in space.

We write the solution in a way similar to that done before[15,19] in the form

$$\Psi(x,t)=\exp\left(-\frac{iv^2t}{2}\right)\sum_s\,[C_s^{(1)}\varphi_s(z_-,t)\times\exp(ivx)+C_s^{(2)}\chi_s(z_+,t)\exp(-ivx)]. \tag{20}$$

Here $z_{\mp}=|x\mp vt|$ and the functions φ_s and χ_s satisfy the equations

$$\left(i\,\frac{\partial}{\partial t}+\frac{1}{2}\,\frac{\partial^2}{\partial x^2}\right)\varphi_s(z,t)=0,\quad \left(i\,\frac{\partial}{\partial t}+\frac{1}{2}\,\frac{\partial^2}{\partial x^2}\right)\chi_s(z,t)=0.$$

Substituting Eq. (20) in Eq. (19) yields the equations

$$\sum_s\left\{C_s^{(1)}\left[\alpha\varphi_s(0,t)+\left.\frac{\partial\varphi_s}{\partial t}\right|_{z=0}\right]+\alpha C_s^{(2)}\chi_s(2vt,t)\times\exp(-2iv^2t)\right\}=0,$$

$$\sum_s\left\{\beta C_s^{(1)}\varphi_s(2vt,t)\exp(-2iv^2t)+C_s^{(2)}\times\left[\beta\chi_s(0,t)+\left.\frac{2\chi_s}{\partial t}\right|_{z=0}\right]\right\}=0. \tag{21}$$

The equations for φ_s and χ_s have solutions of the form

$$\varphi_s=\exp\left(-a_sz+\frac{ia_s^2t}{2}\right),\quad \chi_s=\exp\left(-b_sz+\frac{ib_s^2t}{2}\right),$$

which makes it possible to transform the system of Eqs. (21) for the coefficients $C_s^{(1,2)}$ in the following way:

$$\sum_s\left\{C_s^{(1)}(\alpha-a_s)\exp\left(i\,\frac{a_s^2t}{2}\right)+\alpha C_{s-1}^{(2)}\times\exp\left[\frac{it}{2}\,(b_{s-1}+2iv)^2\right]\right\}=0,$$

$$\sum_s\left\{\beta C_{s-1}^{(1)}\,\exp\left[\frac{it}{2}\,(a_{s-1}+2iv)^2\right]+(\beta-b_s)C_s^{(2)}\times\exp\left(\frac{ib_s^2t}{2}\right)\right\}=0. \tag{22}$$

On going from Eq. (21) to Eq. (22), in the second term of the first equation and in the first term of the second, the summation index s is replaced by $s-1$.

Since Eqs. (22) are satisfied for arbitrary t, the following equalities should be satisfied:

$$a_s=b_{s-1}+2iv,\quad b_s=a_{s-1}+2iv.$$

In addition,

$$\begin{gathered}C_s^{(1)}(\alpha-a_s)+\alpha C_{s-1}^{(2)}=0,\\ \beta C_{s-1}^{(1)}+(\beta-b_s)C_s^{(2)}=0.\end{gathered} \tag{23}$$

Let us set $C_{-1}^{(1)}=C_{-1}^{(2)}=0$, then $a_0=\alpha$ and $b_0=\beta$. The constants $C_0^{(1)}=A$ and $C_0^{(2)}=B$ are not determined by Eqs. (23). From the recurrence relations for a_s and b_s we can obtain

$$a_1=\beta+2iv, \quad b_1=\alpha+2iv,$$
$$a_{2s+1}=\beta+2iv(2s+1), \quad b_{2s+1}=\alpha+2iv(2s+1), \tag{24}$$
$$a_{2s}=\alpha+2iv\cdot 2s, \quad b_{2s}=\beta+2iv\cdot 2s.$$

It also follows from Eq. (23) that

$$C_1^{(1)}=-\frac{\alpha B}{\alpha-\beta-2iv}, \quad C_1^{(2)}=-\frac{\beta A}{\beta-\alpha-2iv}.$$

The recurrence relations (23), as well as the initial values $C_{0,1}^{(1,2)}$ satisfy the following expressions:

$$C_{2l}^{(1)}=\left(-\frac{\alpha\beta}{16v^2}\right)^l \frac{A}{\Gamma(l+1)} \frac{\Gamma(1/2+(\alpha-\beta)/4iv)}{\Gamma(1/2+l+(\alpha-\beta)/4iv)},$$
$$C_{2l+1}^{(1)}=\frac{\alpha}{4iv}\left(-\frac{\alpha\beta}{16v^2}\right)^l \frac{B}{\Gamma(l+1)} \frac{\Gamma(1/2-(\alpha-\beta)/4iv)}{\Gamma(3/2+l-(\alpha-\beta)/4iv)},$$
$$C_{2l}^{(2)}=\left(-\frac{\alpha\beta}{16v^2}\right)^l \frac{B}{\Gamma(l+1)} \frac{\Gamma(1/2-(\alpha-\beta)/4iv)}{\Gamma(1/2+l-(\alpha-\beta)/4iv)},$$
$$C_{2l+1}^{(2)}=\frac{\beta}{4iv^2}\left(-\frac{\alpha\beta}{16v^2}\right)^l \frac{A}{\Gamma(l+1)} \frac{\Gamma(1/2+(\alpha-\beta)/4iv)}{\Gamma(3/2+l+(\alpha-\beta)/4iv)}. \tag{25}$$

The series (20) for the unknown distribution function $\Psi(x,t)$ is convergent for any values of the constants α, β, and x, as well as for arbitrary x, t. We note here that, for $\alpha=\beta$, on using the doubling formulas for the Γ-function one can obtain the same expressions for Ψ as in Ref. 15.

Determining the constants A and B is more complicated, however, when $\alpha\neq\beta$. These constants depend on the parameters α, β, and v and must be determined from the condition that the transition to the case $v=0$ be smooth, when

$$\Psi(x,t)=\sqrt{\frac{\alpha+\beta}{2}}\exp\left[-(\alpha+\beta)\left|x\right|-\frac{i(\alpha+\beta)^2}{2}t\right].$$

It should also be noted that explicit expressions for the ψ-functions in the nonstationary case can also be used for determining the probabilities of changes in the charge state of atomic particles, as has been done before[19] for the case $\alpha=\beta$.

*)E-mail: manko@na.infn.it

---

[1] K. Husimi, Prog. Theor. Phys. **9**, 381 (1953).

[2] J. Schwinger, Phys. Rev. **91**, 728 (1953).

[3] R. Feynman, Phys. Rev. **84**, 108 (1951).

[4] V. P. Ermakov, Univ. Izv. (Univ. Sv. Vladimira, Kiev) **XX**(9), 1 (1880).

[5] H. R. Lewis, Phys. Rev. Lett. **18**, 510 (1966).

[6] H. R. Lewis and W. B. Riesenfeld, J. Math. Phys. **10**, 1458 (1969).

[7] I. A. Malkin, V. I. Man'ko, and D. A. Trifonov, Phys. Rev. D **2**, 1371 (1970).

[8] I. A. Malkin and V. I. Man'ko, Phys. Lett. **32**, 243 (1970).

[9] I. A. Malkin, V. I. Man'ko, and D. A. Trifonov, J. Math. Phys. **14**, 576 (1973).

[10] I. A. Malkin and V. I. Man'ko, Zh. Éksp. Teor. Fiz. **58**, 721 (1970) [Sov. Phys. JETP **31**, 386 (1970)].

[11] C. Groshe, Phys. Lett. A **182**, 28 (1993); Ann. Phys. (N.Y.) **2**, 557 (1993).

[12] M. Moshinsky, Am. J. Phys. **44**, 1037 (1976).

[13] W. Damert, Am. J. Phys. **43**, 531 (1975).

[14] V. V. Dodonov, V. I. Man'ko, and D. E. Nikonov, Phys. Lett. A **162**, 359 (1992).

[15] S. K. Zhdanov and A. S. Chikhachev, Dokl. Akad. Nauk SSSR **218**, 1323 (1974).

[16] A. S. Chikhachev, Zh. Éksp. Teor. Fiz. **107**, 1153 (1995) [JETP **80**, 641 (1995)].

[17] N. M. Nussenzweig, in *Symmetries in Physics (Proc. of the I International Symposium Held in Honor of Prof. M. Moshinsky*, 1991, A. Frank and K. B. Wolf (eds.), Springer-Verlag, Berlin, Heidelberg (1992).

[18] I. A. Malkin and V. I. Man'ko, *Dynamic Symmetries and Coherent States of Quantum Systems* [in Russian], Nauka, Moscow (1979).

[19] W. Dappen, J. Phys. B **10**, 2399 (1977).

Translated by D. H. McNeill

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Oscillatory disintegration of nonevolutionary magnetohydrodynamic discontinuities

S. A. Markovskii

Sternberg Astronomical Institute, Moscow State University, 119899 Moscow, Russia

(Submitted 28 April 1997)
Zh. Éksp. Teor. Fiz. **113**, 615–628 (February 1998)

Trans-Alfvénic shock waves are considered in the approximation of small amplitude and almost parallel propagation of the magnetic field. Such shocks are nonevolutionary, since the problem of time evolution of their small perturbation does not have a unique solution. Therefore, they cannot exist as stationary configurations and must disintegrate or transform to some more general, nonsteady flow. The disintegration configuration necessarily includes an Alfvén discontinuity that is also nonevolutionary. It is shown that the contradiction inherent in the nonevolutionary configuration is removed if its time evolution has the form of oscillatory disintegration, i.e., reversible transformation of one type of discontinuity to the other. In this process fast and slow shock or rarefaction waves as well as contact discontinuities are emitted. © *1998 American Institute of Physics.* [S1063-7761(98)01802-2]

1. INTRODUCTION

The problem of disintegration of hydrodynamic discontinuities has a long history since the publication of the paper by Kotchine.[1] He considered the disintegration of an arbitrary discontinuity into a set of other discontinuities and rarefaction waves in the framework of nonmagnetic hydrodynamics. Some time later, Bethe[2] studied the disintegration of a shock wave. Magnetic field complicates the situation, enlarging the number of possible disintegration configurations. For a small-amplitude arbitrary discontinuity such configurations were obtained by Lyubarskii and Polovin.[3] In general, the problem cannot be solved in an analytic form. Gogosov[4] has given a quantitative solution that determines the type of the configuration, depending on the flow parameters.

The disintegration of a shock wave is closely related to the problem of its evolutionarity, formulated in Refs. 5–7. It is suggested that small perturbations should be imposed on the discontinuity surface to study the question of its disintegration. In this case small-amplitude waves occur on both sides of the surface. The amplitudes of these waves are related by the linearized boundary conditions obtained from conservation laws at the discontinuity. If the amplitudes of the outgoing waves cannot be determined unambiguously from these conditions by the amplitudes of the incident waves, then the problem of the time evolution of the infinitesimal perturbations does not have a unique solution, and the discontinuity is called nonevolutionary. This problem is encountered when the number of unknown parameters (the amplitudes of the outgoing waves and the discontinuity displacement) is incompatible with the number of independent equations.

Since a physical problem must have a unique solution, it is not correct to assume that the perturbation of a nonevolutionary discontinuity is infinitesimal. Such a discontinuity cannot exist in a real medium as a stationary configuration, because the infinitesimal perturbation leads to a finite variation of the initial flow. This variation is the disintegration of the discontinuity into other discontinuities, which move away from the place of their formation, or a transition to a more general nonsteady flow. In the ideal medium the disintegration is instantaneous in the sense that the secondary discontinuities become separated in the beginning of the disintegration process. In a dissipative medium the spatial profiles of the magnetohydrodynamic (MHD) properties are continuous. Nevertheless, the principal result remains the same, the flow is rearranged toward a nonsteady state, and after a large enough period of time the disintegration manifests itself.

The evolutionarity requirement gives additional (compared to the Zemplen theorem) restrictions on the flow parameters at the shock surface. They follow from the fact that the direction of wave propagation (toward the discontinuity surface or away from it) and hence the number of the outgoing waves depends on the flow velocity at the surface. If it is large enough, then the given wave may be carried down by the flow. Therefore, at an evolutionary discontinuity the flow velocity must be such that it provides the compatibility of the set of boundary equations. This form of evolutionarity condition was applied to MHD shock waves in Refs. 8–10.

As a result, the fast (I→II) and slow (III→IV) shocks, for which the flow velocity both upstream and downstream is larger and smaller than the Alfvén velocity, respectively, are evolutionary, while the trans-Alfvénic shock waves (TASWs) are not. Here the Roman numbers indicate the states upstream and downstream of the shock, in which the values of the normal flow velocity fall into the intervals separated by the three phase velocities: fast magnetosonic, Alfvén, and slow magnetosonic velocities. These states are arranged in order of increasing entropy.

The important fact that favors the nonexistence of nonevolutionary shock transitions is that they can (while the evolutionary ones cannot) be realized also through a set of several shock and rarefaction waves.[11–13] One more argument for the nonexistence of nonevolutionary shocks is that they are isolated solutions of the Rankine–Hugoniot problem which do not have neighboring solutions corresponding to small deviation of boundary states.[14] This is confirmed by the fact that the configurations neighboring such shocks are time dependent.[15,16] For such configurations the coplanarity of the boundary states is violated and therefore they are not solutions of the Rankine–Hugoniot problem.

The problem of structural instability of MHD shocks has recently progressed to a new point due to the consideration of a nonplanar shock structure.[17–20] Kennel *et al.*[19] discussed nonplanar shocks of small amplitude. They have demonstrated that the structure of nonevolutionary, TASWs, is not unique. Namely, the transitions II→III can be connected by two integral curves, left-hand and right-hand polarized, and the transitions I→III and II→IV allow an infinite number of connecting integral curves. These conclusions are in agreement with those of Hau and Sonnerup,[18] who analyzed the stationary points of the MHD equations corresponding to the boundary states of the shock transitions in the case where the magnetic diffusivity is the only nonzero transport coefficient. Recall that under the assumption of a planar shock structure the trans-Alfvénic transitions also do not have a unique structure for all values of the dissipative transport coefficients.[21–23]

It can be shown[19] that the integral over the shock thickness of the out-of-plane component of the magnetic field is independent of time if the upstream and downstream states of the small amplitude shock satisfy the Rankine–Hugoniot conditions. This means that the flow outside the shock is in one plane. This integral, which characterizes the nonplanar structure, remains constant during the evolution of the initial profile, and it labels uniquely the integral curve that connects the given states. For the evolutionary, I→II and III→IV, shocks the flow is planar, and the integral is zero. For the shock transition II→III the integral takes two values with equal modules and opposite signs, and for the transitions II→IV and I→III it falls into some interval which describes a one-parameter family of structures.

To remove the ambiguity of the solution for the TASWs, Kennel *et al.*[19] postulated that in addition to the boundary states, the integral characterizing the nonplanar structure should be fixed. The solution of the boundary value problem describing the shock will then be unique. To assure this uniqueness, however, one must assume that the configuration contains one shock. This is not the only possibility. As discussed above, the trans-Alfvénic shock transitions can be realized also through a set of several shock and rarefaction waves. Thus, the ambiguity of the structure connecting the nonevolutionary boundary states is not lifted,[24] and the TASW is structurally unstable. Nevertheless, the conservation of the quantity that fixes the structure of the nonplanar shock is an additional factor that governs the disintegration process.

Indeed, at a TASW the tangential magnetic field changes sign. Consequently, it must change sign at a secondary discontinuity. This may take place either at another TASW or at an Alfvén discontinuity. As is known,[25] the Alfvén discontinuity is also nonevolutionary in the presence of arbitrary, small, but nonzero dissipation. Since the evolution of a nonplanar shock is related to its structure, it cannot be assumed that the dissipation is absent. Therefore, the Alfvén discontinuity also cannot exist as a stationary configuration. In the present paper we suggest a new scenario for the evolution of the TASW, oscillatory disintegration, i.e., reversible transformation of the Alfvén discontinuity. Such a form of evolution resolves the contradiction inherent in the nonevolutionary configurations.

We consider the small-amplitude shocks that propagate almost parallel to the magnetic field. This approximation allows us to use an analytical approach, and, at the same time, to determine some features of the behavior of finite-amplitude discontinuities. In Sec. 2 we obtain the disintegration configuration for the TASW of small amplitude that propagate almost parallel to the magnetic field. In Sec. 3 we discuss the evolutionarity and some other properties of nonplanar shocks. In Sec. 4 we describe the time evolution of the TASW in the case where the transverse magnetic field is not small and in the case of almost parallel propagation. Finally, we present our conclusions in Sec. 5.

2. STRUCTURAL INSTABILITY OF TRANS-ALFVÉNIC SHOCKS

We first consider the disintegration configurations of small-amplitude, almost parallel MHD shocks. We choose the frame of reference in which $\mathbf{B}\|\mathbf{v}$ and the x axis is directed along the normal to the discontinuity. We proceed from the following jump conditions at the discontinuity surface:

$$\Delta(\rho v_x)=0, \tag{1}$$

$$\Delta p+\rho v_x\Delta v_x+\frac{1}{8\pi}\,\Delta(B_y^2)=0, \tag{2}$$

$$\rho v_x\Delta v_y-\frac{B_x}{4\pi}\,\Delta B_y=0, \tag{3}$$

$$v_y=v_xB_y/B_x, \tag{4}$$

$$\frac{1}{2}\,\rho v_x\Delta(v^2)+\frac{\gamma}{\gamma-1}\,\Delta(pv_x)=0. \tag{5}$$

In solving this set of equations we imply that the variations Δ of all MHD properties, except B_y, are small compared to their values at a reference state, and that $B_y\ll B_x$. To the lowest order in the small parameters B_y/B_x and $\Delta\rho/\rho$ Eqs. (1)–(5) have the following solutions which describe the relationship between the jumps of the MHD properties. For the first four solutions $\Delta B_y\gtrsim B_y$.

$$\Delta B_y=-B_y\left(1\pm\sqrt{1+2\,\frac{V_{Ax}^2-V_s^2}{V_{Ay}^2}\,\frac{\Delta\rho}{\rho}}\right)\equiv B_yA, \tag{6}$$

$$\Delta v_y=-\varepsilon V_{Ay}A, \tag{7}$$

$$\Delta v_x=\varepsilon V\Delta\rho/\rho, \tag{8}$$

$$\Delta p=V_s^2\Delta\rho, \tag{9}$$

where $V=V_{Ax}$ and $\mathbf{V}_A=\mathbf{B}/\sqrt{4\pi\rho}$ is the Alfvén velocity. Here $\varepsilon=+1$ for the waves moving in the positive x direction and $\varepsilon=-1$ for the waves moving in the opposite direction. We assume for definiteness that $V_{Ax}>V_s$. The plus

(«+») sign in Eq. (6) will then correspond to a TASW and the minus («−») sign will correspond to a fast shock wave. In zeroth approximation the propagation velocity of both waves is V_{Ax}. The TASW is of the II→III type, i.e.,

$$V_{+1}>v_{x1}>V_{Ax1}, \quad V_{Ax2}>v_{x2}>V_{-2}$$

if

$$-1\leqslant\frac{B_{y1}}{B_{y2}}\leqslant-\frac{1}{2}. \tag{10}$$

and of I→III type, i.e.,

$$v_{x1}>V_{+1}, \quad V_{Ax2}>v_{x2}>V_{-2}$$

if

$$-\frac{1}{2}\leqslant\frac{B_{y1}}{B_{y2}}\leqslant 0. \tag{11}$$

Here

$$V_{\pm}^2=\frac{1}{2}\left[V_A^2+V_s^2\pm\sqrt{(V_A^2+V_s^2)^2-4V_s^2V_{Ax}^2}\right]$$

are the phase velocities of the fast (+) and slow (−) small-amplitude waves, and the subscripts «1» and «2» indicate the states upstream and downstream of the shock, respectively.

Two more solutions of the set of equations (1)–(5), for which $\Delta B_y\ll B_y$, correspond to slow shocks. They are given by the formulas

$$\Delta B_y=-\frac{V_s^2B_y}{V_{Ax}^2-V_s^2}\,\frac{\Delta\rho}{\rho}, \tag{12}$$

$$\Delta v_y=\varepsilon\,\frac{V_sV_{Ay}V_{Ax}}{V_{Ax}^2-V_s^2}\,\frac{\Delta\rho}{\rho}, \tag{13}$$

and by Eqs. (8) and (9) with $V=V_s$, where $V_s=\sqrt{\gamma p/\rho}$ is the sound velocity. Expressions (8), (9), (12), and (13) coincide with those for small-amplitude waves. At the Alfvén discontinuity

$$\Delta B_y=B_y(\pm 1-1), \quad \Delta v_y=-\varepsilon V_{Ay}(\pm 1-1), \tag{14}$$

where the plus (+) sign is used if the discontinuity is absent and the minus (−) sign is used if it rotates the magnetic field through 180°. Finally, the only nonzero jump at a contact discontinuity is $\Delta\rho$.

Assume now that the TASW with the amplitude $\Delta_0\rho$ moves in the positive x direction. If the shock transition can be represented as a set of more than one discontinuity, the amplitudes of the secondary discontinuities are determined from the condition that the sums of the jumps of the MHD properties at them are equal to those at the initial shock. It should be mentioned that in so doing the variation of B_y at the secondary waves must be taken into account in Eqs. (6)–(9) and (12)–(14), while the other quantities equal to their values upstream of the initial shock may be substituted. We thus find that in zeroth approximation the trans-Alfvénic shock transition can be realized through a fast shock, with the same amplitude $\Delta_f^{(+1)}\rho=\Delta_0\rho$, and the Alfvén discontinuity moving in the same direction. However, since these secondary waves move with zero velocity with respect to each other, there is no disintegration.

Let us now solve Eqs. (1)–(5), taking into account higher-order terms. For simplicity we assume that $B_y^2/B_x^2 \lesssim\Delta\rho/\rho$. Otherwise (when $B_y^2/B_x^2\gg\Delta\rho/\rho$), the approximation of almost parallel propagation is violated. The corrections to the quantities (6)–(9) are determined by the expressions that follow from the expansion of Eqs. (1)–(5) in the small parameters B_y/B_x and $\Delta\rho/\rho$.

$$\Delta B_y=B_y(A+a), \tag{15}$$

$$\Delta v_y=-\varepsilon V_{Ay}\left[A+a-\frac{1}{2}\,(1+A)\,\frac{\Delta\rho}{\rho}\right]. \tag{16}$$

$$\Delta v_x=\varepsilon V_{Ax}\,\frac{\Delta\rho}{\rho}\left[1-\frac{1}{2}\left(1-\frac{1}{A}\right)\frac{\Delta\rho}{\rho}\right], \tag{17}$$

$$\Delta p=V_s^2\Delta\rho+\frac{V_{Ax}^2(\Delta\rho)^2}{\rho A}-\frac{B_y^2}{4\pi}\,a(1+A). \tag{18}$$

Here

$$a=\frac{2A^3+(\gamma+4)A^2+b_1A+b_2}{2A(1+A)}\,\frac{\Delta\rho}{\rho}, \tag{19}$$

$$b_1=2+2b_2-(\lambda+3)\,\frac{V_{Ax}^2}{V_{Ay}^2}\,\frac{\Delta\rho}{\rho}, \quad b_2=2\,\frac{V_s^2}{V_{Ay}^2}\,\frac{\Delta\rho}{\rho}. \tag{20}$$

The velocity in front of such shocks is

$$v_x=-\varepsilon V_{Ax}\left[1+\frac{1}{2}\left(1+\frac{1}{A}\right)\frac{\Delta\rho}{\rho}\right]. \tag{21}$$

Next, we set equal again the sums of the jumps at the secondary waves to those at the initial TASW. In substituting the relationship between the MHD properties for the waves absent in zeroth approximation it is sufficient to use Eqs. (6)–(9), (12), and (13), which are valid in the lowest order, because the amplitudes of these waves are small compared to the amplitude of the initial shock.

As a result, we find that the fast wave moving in the direction opposite to the initial shock has the amplitude

$$\Delta_f^{(-1)}\rho=\frac{1}{2}\,\frac{V_{Ay1}^2A_-(1+A_-)\Delta_0\rho}{V_{Ax}^2-V_s^2}. \tag{22}$$

The amplitude of the contact discontinuity is

$$\Delta_c\rho=-(\gamma-1)\Delta_0\rho\,\frac{V_{Ay1}^2}{V_s^2}\,\sqrt{1+2\,\frac{V_{Ax}^2-V_s^2}{V_{Ay1}^2}\,\frac{\Delta_0\rho}{\rho}}, \tag{23}$$

and the amplitudes of the slow waves moving in the positive and negative x directions are

$$\Delta_s^{(\varepsilon)}\rho=-\frac{1}{2}\,\Delta_c\rho+\varepsilon\,\frac{V_{Ax}}{V_s}\,\Delta_f^{(-1)}\rho+\frac{\rho(1+A_-)V_{Ay1}^2}{2(V_{Ax}^2-V_s^2)}$$
$$\times\left[a_++a_-+\varepsilon\,\frac{V_{Ax}}{V_s}\left(a_++a_--\frac{1}{2}\,\frac{\Delta_0\rho}{\rho}\right)\right]. \tag{24}$$

Here A and a are given by Eqs. (6), (19), and (20), in which $B_y=B_{y1}$ and $\Delta\rho=\Delta_0\rho$, and the subscripts ‘‘+’’ and ‘‘−’’

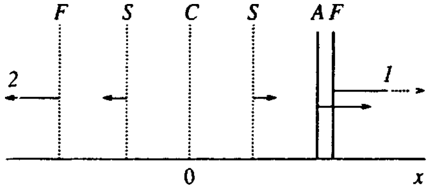


FIG. 1. Disintegration configuration for the trans-Alfvénic shock in the case $V_s \gg V_{Ay}$. In zeroth approximation the dotted lines are absent.

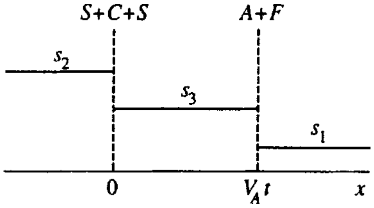


FIG. 2. Time-dependent entropy profile after the disintegration of the trans-Alfvénic shock in case $V_s \lesssim V_{Ay}$.

correspond to the sign in Eq. (6). It can be shown, with the help of Eqs. (17) and (21), that the absolute value of the normal flow velocity behind the fast shock moving in the positive x direction is larger than the normal Alfvén velocity in front of the Alfvén discontinuity. Hence, these waves become separated as time goes on (Fig. 1).

Note that if V_s is comparable with V_{Ax}, then the flow in the fast and trans-Alfvénic shocks, which is determined by Eqs. (6)–(9), is isentropic, but only to the lowest order. The jump of the entropy at such shocks is

$$\Delta s = \frac{\gamma}{4} \frac{V_{Ay}^2}{V_s^2} \frac{\Delta\rho}{\rho} A^2. \tag{25}$$

In case where $\Delta\rho/\rho \ll V_{Ay}^2/V_{Ax}^2$ this expression coincides with that obtained by Bazer and Ericson.[26] After the disintegration, the difference between the entropy jumps at the trans-Alfvénic and the fast shock is taken by the contact discontinuity at which the entropy jump is of the same order of magnitude as the density jump

$$\Delta s = \left(\frac{\partial s}{\partial \rho}\right)_p \Delta\rho = -\frac{\gamma}{\gamma - 1} \frac{\Delta\rho}{\rho}. \tag{26}$$

Thus, the consideration of the corrections to zeroth approximation reveals two properties of the disintegration configuration. First, the fast shock and the Alfvén discontinuity moving in the same direction as the initial shock acquire a small relative velocity. Second, the amplitudes of the discontinuities, which have finite velocity with respect to each other, become nonzero.

We emphasize that the absence of disintegration in zeroth approximation is essentially a consequence of the assumption that the tangential magnetic field is small. This result is consistent with the fact that the exactly parallel TASW can be represented as a switch-on and a switch-off shock, but they do not become separated.[10,11] However, the parallel shock disintegrates when it collides with small-amplitude shocks incident on both sides of the discontinuity surface.[11] It should be mentioned that Eqs. (6) and (7) are not valid for the exactly parallel shock, because it necessarily has a finite amplitude if V_s is not close to V_{Ax}.

We can also consider the case $V_s \ll V_{Ax}$, which was discussed by Kennel *et al.*[19] Note that, on the other hand, V_s^2 must be much larger than $V_{Ay}^2 \Delta\rho/\rho$ (or V_{Ay}^4/V_{Ax}^2) for a small-amplitude shock. The character of the disintegration changes significantly when $V_s \lesssim V_{Ay}$. Under this condition the flow is not isentropic in the lowest order [see Eqs. (23), (25), and (26)]. To this order the jumps $\Delta_c s$, $\Delta_c \rho$, and $\Delta_s^{(\varepsilon)} \rho$, are not equal to zero, in contrast with the case $V_s \gg V_{Ay}$. At the same time, the velocity of the slow shocks V_s is small. Therefore, in zeroth approximation they are not separated from the contact discontinuity which is at rest with respect to the medium.

As a result, the initial TASW disintegrates into two structures. The first structure, denoted by $A+F$, is formed by the Alfvén discontinuity and the fast shock, which are at rest with respect to each other. The second, denoted by $S+C+S$, is formed by two slow shocks and the contact discontinuity. These structures have the finite relative velocity V_{Ax}. The peculiarity of such a configuration is that the only nonzero total jump at the structure $S+C+S$ is $\Delta_c s$, because, as can be seen from Eqs. (22)–(24), the total density jump equals zero when $V_s \lesssim V_{Ay}$. Consequently, the disintegration in this case takes place in the lowest order, although its only manifestation is the time-dependent entropy profile along the x axis (Fig. 2). In higher orders the profiles of the other MHD properties also become nonsteady.

3. EVOLUTIONARITY OF NONPLANAR SHOCKS

Next, we consider the influence of the nonplanar structure on the shock evolution. If the boundary states of the shock are coplanar, the evolution of the configuration is characterized by additional conservation laws. This can be understood, for example, from the z-component of the induction equation,

$$\frac{\partial B_z}{\partial t} = -\frac{\partial}{\partial x}\left(B_z v_x - B_x v_z - \nu_m \frac{\partial B_z}{\partial x}\right), \tag{27}$$

where ν_m is the magnetic diffusivity. Since the flow outside the discontinuity is planar and homogeneous, the integration over the thickness of the transition layer $\Delta x = x_2 - x_1$ yields

$$\frac{\partial}{\partial t} \int_{x_1}^{x_2} B_z dx = 0. \tag{28}$$

Similarly, it follows from the z-component of the momentum equation,

$$\frac{\partial \rho v_z}{\partial t} = -\frac{\partial}{\partial x}\left(\rho v_x v_z - \frac{1}{4\pi} B_x B_z - \eta \frac{\partial v_z}{\partial x}\right), \tag{29}$$

that the integral of the quantity ρv_z is also conserved in the process of time evolution of the configuration,

$$\frac{\partial}{\partial t} \int_{x_1}^{x_2} \rho v_z dx = 0. \tag{30}$$

Here η is the shear viscosity.

As shown in Ref. 19, if the integral of B_z over the profile of the small amplitude almost parallel II→III shock is prescribed, then the ratio B_{y1}/B_{y2} of the upstream to the downstream values of the transverse magnetic field is fixed. Thus, the state behind the shock is unambiguously determined by the state in front of it. This gives us an additional equation for the amplitudes of the waves that occur after the disintegration. Assuming that this equation is valid to the second order in the amplitude, it can be shown with the help of Eqs. (12), (13), and (15)–(18) that the amplitudes of the initial and secondary TASWs are equal, while all other secondary waves are absent. The same is true for the I→III shock. The difference is that in the latter case there is no additional equation, because the given integral over the profile allows infinite number of boundary states.[19] However, the fast wave moving in the direction of the initial shock is absent. Therefore, the number of equations is again equal to the number of secondary waves. These equations have only a trivial solution. Consequently, in the approximation under consideration a TASW cannot appear after the disintegration as a secondary wave.

It should be mentioned that the conclusion about the relationship between the integral quantity, which characterizes the structure, and the amplitude of the wave is made in Ref. 19 only for isentropic flows. As shown in Sec. 2, this is not always the case for the almost parallel propagation. Nevertheless, this conclusion may be also made on the basis of the following reasoning. The consideration of the stationary points of the MHD equations, corresponding to the boundary states of the shock transitions, shows that the transition II→III can be realized through two integral curves, right-hand and left-hand polarized, in contrast with the transitions I→III (II→IV) and I→IV which are described by one- and two-parameter families of curves, respectively (see, e.g., Ref. 18). Since the structure of the II→III shock does not contain free parameters other than the amplitude (although it is nonunique), for the given quantity characterizing the structure and the given state in front of the shock its amplitude is fixed.

Below we restrict the discussion to the shocks of the II→III type. The evolution of these shocks is the most important feature for the following reason. In contrast with the other types, the nonevolutionarity of the II→III shocks is essentially based on the fact that for normally propagating waves the equations for the Alfvén perturbations are separated from those for the magnetosonic and entropy perturbations.[10] Therefore, the evolutionary criterion must be satisfied separately for both groups of waves. This makes the shock nonevolutionary, although the total number of perturbations is compatible with the total number of boundary conditions. Since under the assumption of a nonplanar shock structure the separation does not take place,[20] this argument formally does not hold. Nevertheless, the coupling of Alfvén modes to magnetosonic and entropy modes does not alter the conclusions made on the basis of the evolutionarity principle. Let us clarify this point.

In the linear approximation Eqs. (27) and (29) for the perturbations proportional to $\exp(i\omega t)$ take the following form after the integration over x:

$$i\omega\int_{x_1}^{x_2}\delta B_z dx=-\Delta\left(v_x\delta B_z-B_x\delta v_z-\nu_m\frac{\partial\delta B_z}{\partial x}\right), \tag{31}$$

$$i\omega\int_{x_1}^{x_2}v_z\delta\rho dx+i\omega\int_{x_1}^{x_2}\rho\delta v_z dx=-\Delta\left(\rho v_x\delta v_z-\frac{1}{4\pi}B_x\delta B_z-\eta\frac{\partial\delta v_z}{\partial x}\right), \tag{32}$$

where δ is the small perturbation, and the unperturbed quantities correspond to the stationary shock. The term responsible for the coupling is the first integral on the left-hand side of Eq. (32).

Let us assume now that the Alfvén wave, which transfers only the perturbations δv_z and δB_z, is incident on the II→III shock. In this case there is one outgoing Alfvén wave, whose amplitude is the unknown parameter that should be determined from Eqs. (31) and (32). Because the perturbation $\delta\rho$ enters into Eq. (32), when $v_z\neq 0$, the latter becomes an additional equation for the amplitudes of outgoing magnetosonic and entropy waves which may be generated by the incident Alfvén wave. The perturbation $\delta\rho$ inside the transition layer depends on the amplitudes of the waves outside it and on the stationary shock structure. Since $\delta\rho$ should be determined from Eq. (32), the term with $\delta\rho$ cannot be much smaller than all other terms. Estimating the first terms on the left- and right-hand sides of Eq. (32), we obtain in the order of magnitude

$$\omega v_{z0}\delta\rho_0\Delta x\sim\rho_0 v_{x0}\Delta\delta v_z, \tag{33}$$

where the subscript 0 indicates some characteristic values.

In general, $\Delta\delta v_z\sim\delta v_{z0}$. In addition, the unperturbed quantities and the discontinuity thickness do not depend on the frequency of the perturbation. Consequently, for the perturbation with small enough ω the quantity $\delta\rho_0/\rho_0$ is arbitrarily large compared to $\delta v_{z0}/v_{z0}$. However, this result is valid for any nonevolutionary shock; namely, as discussed in Sec. 1, infinitesimal incident perturbation (δv_z in the present case) causes a finite variation of the flow ($\delta\rho$). Therefore, since the coupling of the modes is weak, the contradiction inherent in the II→III shock is not resolved even in the absence of the separation of the boundary conditions.

In connection with the nonunique shock structure, Wu[20] argued that the shocks of the remaining types become evolutionary if the free parameters characterizing their structure are added to the total number of perturbations of the shock. However, the additional free parameters on their own also do not resolve the contradiction. In the presence of dissipation such parameters are provided, in particular, by the amplitudes of purely dissipative waves, which are absent in the ideal medium.[25] The problem is that the boundary conditions, that follow from the conservation laws at the shock, are incompatible. Therefore, a free parameter contributes to the evolution only if it enters into the conservation laws. For example, purely dissipative waves damp within the length of the order of the shock thickness. Consequently, their amplitudes do not enter into the boundary conditions, and they should be disregarded when solving the problem of evolution of a shock wave, unless it is of switch-off or switch-on type.[25] At the same time, it can be shown that in the case of dissipative discontinuities inside inviscid shock waves the additional dissipative waves affect the evolution condition. Apparently, the structure variations are also confined within the transition layer.

4. TIME EVOLUTION OF NONEVOLUTIONARY DISCONTINUITIES

Let us now turn to the time evolution of the TASWs. For an illustration we first consider the case in which the ratio B_y/B_x is not small. To the first order in $\Delta\rho/\rho$ we can then reduce the expressions for the jumps of the MHD properties at the TASW as follows:[12]

$$\Delta_0 Q_j = \Delta_A Q_j + A_{0j}\Delta_0\rho. \tag{34}$$

Here $\mathbf{Q}=(\rho,p,v_x,v_y,B_y)$ is the vector of state, i.e., the set of MHD properties, A_{0j} are the known coefficients, and Δ_A are the jumps at the Alfvén discontinuity given by Eq. (14). In the present case the jump $\Delta|B_y|$ is small compared to $|B_y|$; hence, the inequality (10) is satisfied, and the TASW is of the II→III type.

The jumps at the fast and slow shock (or rarefaction) waves coincide with those at the corresponding small-amplitude waves. As a result, the equations that determine the disintegration configuration take the form

$$\sum_i A_{ij}\Delta_i\rho + \Delta_A Q_j = A_{0j}\Delta_0\rho + \Delta_A Q_j, \tag{35}$$

where the subscript i indicates the type of the discontinuity, and A_{ij} are known quantities: The solution of these algebraic equations

$$\Delta_i\rho = a_i\Delta_0\rho \tag{36}$$

expresses the amplitudes of secondary waves in terms of the amplitude of the initial TASW. The explicit expressions for the quantities a_i are given in Ref. 12.

Thus, the initial nonevolutionary shock may disintegrate into a contact discontinuity, magnetosonic waves, and an Alfvén discontinuity. At the same time, as shown by Roikhvarger and Syrovatskii,[25] the Alfvén and the contact discontinuities are also nonevolutionary in the presence of an arbitrarily small but nonzero dissipation and heat conduction. This stems from the fact that the flow velocity in this case is equal to the phase velocity of the Alfvén and the entropy wave, respectively. As a result, the wavelength of the small perturbation with a fixed frequency tends to zero. This makes it necessary to account for the dissipation and, as a consequence, leads to the occurrence of additional outgoing (dissipative) waves that damp within the length much larger than the discontinuity thickness.

However, the nonevolutionarity does not lead to the disintegration of the discontinuity if it cannot be represented as a set of several discontinuities. This is the case for a contact discontinuity. Indeed, under the condition that B_y is not small the jumps Δv_y and ΔB_y at a TASW or at an Alfvén discontinuity are much larger than all other jumps. To provide no field reversal in sum, the disintegration configuration must contain two discontinuities at which B_y and v_y change sign or it must not contain them at all. In the latter case the equations for the jumps have only the trivial solution. It can be readily verified, with the help of Eqs. (14) and (35), that in the former case Δv_y and ΔB_y cannot be compensated simultaneously, since Δv_y depends on the direction of propagation, while ΔB_y does not. Hence, the contact discontinuity is structurally stable. In contrast, the Alfvén discontinuity is unstable, and it may disintegrate.

On this basis we suggest a new scenario for evolution of the configurations with the magnetic field reversal, oscillatory disintegration. Under the action of an infinitesimal perturbation the TASW disintegrate into a system of waves including the Alfvén discontinuity. In this process the integrals (28) and (30) over the nonplanar profile of the initial wave are conserved and are equal to the integrals over the profile of the Alfvén discontinuity, while the flow in the remaining secondary waves is plane. In contrast with a shock wave, the out-of-plane structure of an Alfvén discontinuity is not related to the boundary states due to their degeneration. Therefore, such a disintegration configuration can always be adjusted to the initial discontinuity. The Alfvén discontinuity, in turn, also disintegrates, producing the TASW with an amplitude equal to that of the initial wave, which is unambiguously determined by the quantity fixing the structure.

The amplitudes of the remaining waves satisfy Eq. (35), in which $\Delta_0\rho$ is replaced on the right-hand side. Therefore, they are given by Eq. (36), in which a_i is replaced by $-a_i$; i.e., shock waves instead of rarefaction waves occur and vice versa, compared to the case of disintegration of the TASW. After that the process is repeated. The waves of different types may catch up with and outrun each other during their propagation. Since the waves are structurally stable, their types do not change in this process. If the characteristic time between the disintegrations is not small, the waves of the same types catch up with each other at infinite time. In the approximation of small amplitudes the secondary waves do not interact with each other. This means that the feedback effect of the disintegration on the nonevolutionary discontinuity manifests itself in higher orders.

It should be mentioned that the contradiction associated with the nonevolutionary behavior of the Alfvén discontinuity is also resolved if it has a time-dependent thickness. To some extent, the situation is similar to that for the corrugationally unstable shocks in the nonmagnetic hydrodynamics.[27,28] As is known,[29,30] such shock transitions can always be represented as a combination of several discontinuities. This makes it possible to assume that the unstable shocks may disintegrate rather than undergo the growing undulation.[30] However, as in the present case, the physical mechanism that distinguishes whether the discontinuity maintains itself during the evolution or transforms to another configuration remains unclear.

Let us return to the almost parallel shocks. In this case the equations that determine the disintegration configuration are not linear equations, in contrast with Eq. (35). Nevertheless, to the lowest order, the amplitudes of waves occurring after the disintegration of the Alfvén discontinuity and the TASW are also equal in absolute value and have opposite signs. This can be shown by indicating that the expressions

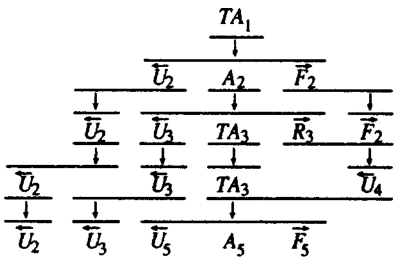


FIG. 3. Scheme of oscillatory disintegration of the initial trans-Alfvénic shock (denoted by TA) into the Alfvén discontinuity (A), the fast shock (F) and rarefaction (R) waves, and the structure $U=S+C+S$. The vertical arrows show the time evolution and the horizontal arrows indicate the velocity with respect to the nonevolutionary discontinuity.

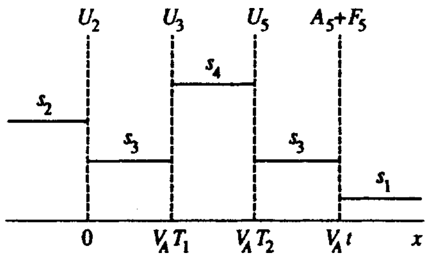


FIG. 4. Entropy profile in the state of the system after the complete cycle of the disintegration. T_1 and T_2 are the moments of time at which the disintegration takes place.

for the variations of the MHD properties Δ at the rarefaction waves coincide with those for shock waves. In the approximation of small B_y the equations describing the rarefaction waves (see, e.g., Ref. 31) take the form

$$\rho\,\frac{dB_y^2}{d\rho}=B^2-4\pi\rho V_s^2\pm\left(|B_x^2-4\pi\rho V_s^2|\right.$$
$$\left.+\frac{(B_x^2+4\pi\rho V_s^2)B_y^2}{|B_x^2-4\pi\rho V_s^2|}\right),\tag{37}$$

$$\frac{dv_y}{dB_y}=-\frac{1}{\sqrt{4\pi\rho}},\tag{38}$$

$$\frac{dv_x}{d\rho}=\frac{V_A}{\rho},\tag{39}$$

$$\frac{dp}{d\rho}=V_s^2.\tag{40}$$

We thus obtain to the lowest order Eqs. (6)–(9), (12), and (13) for the differences Δ of the downstream to the upstream values.

In the case of small B_y the difference between the velocity of the fast shock and rarefaction waves is small. Therefore, if the disintegration takes place at finite intervals, these waves may catch up with each other at a finite time. Moreover, under the condition $V_s \lesssim V_{Ay}$, the feedback effect of the disintegration on the initial wave is of the same order of magnitude as its amplitude. A disintegration scheme in this case is presented in Fig. 3. In this figure the lowest-order waves only are shown, $U=S+C+S$, and the TASW are denoted by TA. As can be seen from Fig. 3, after the fast rarefaction R_3 catches up with the fast shock F_2 the configuration U_4 that moves toward TA_3 remains there. Since the TASW are nonevolutionary and structurally unstable, their interaction results in disintegration.

After the complete cycle of the oscillatory disintegration the system comes to the state shown in Fig. 4. The quantities s_3 and s_4 are

$$s_3=s_1+\frac{\gamma}{4}\,\frac{V_{Ay1}^2}{V_s^2}\,\frac{\Delta_0\rho}{\rho}\left(1-\sqrt{1-\frac{V_{Ax}^2}{V_{Ay1}^2}}\right)^2,\tag{41}$$

$$s_4=s_1+\gamma\frac{V_{Ay1}^2}{V_s^2}\,\frac{\Delta_0\rho}{\rho}\left(1+\frac{V_{Ax}^2}{V_{Ay1}^2}\right).\tag{42}$$

This state resembles the state shown in Fig. 2, except for the configurations U_3 and U_5 that compensate each other. Thus, the nonevolutionary shock emits the discontinuities $S+C+S$ in the process of its evolution.

5. CONCLUSIONS

We have examined the disintegration of small-amplitude nonevolutionary shock waves. We have shown that in case of almost parallel propagation to the magnetic field the shock is structurally unstable in the second order in its amplitude. Such a shock transition can be represented as a set of several discontinuities moving with respect to each other. As a result, the shock structure is ambiguous not only because the boundary states are connected by a nonunique integral curve, but also because the shock transition can be realized through the single shock and through the configuration that consists of more than one discontinuity.

However, the disintegration configuration necessarily includes an Alfvén discontinuity that is also nonevolutionary. The contradiction can be resolved if the further time evolution has the form of oscillatory disintegration, i.e., reversible transformation to the Alfvén discontinuity. In this process shock and rarefraction waves, as well as contact discontinuities, which move with respect to each other, are emitted.

Such a process is similar to spontaneous emission of small-amplitude waves by a shock wave without a magnetic field. This phenomenon was observed in the laboratory experiments.[32,33] It appears in the special case of corrugational instability of the shock when its small perturbation does not grow with time, but propagates away from the discontinuity surface in the form of nondamping waves, whose energy is supplied from the whole moving medium. The similarity is natural, because in this case the reflection and refraction coefficients at the discontinuity surface tend to infinity in the presence of incident waves (see, e.g., Ref. 34). Consequently, such a shock is nonevolutionary, since the problem of small perturbation does not have a solution.[24]

At the same time, the oscillatory disintegration has two significant distinctions. First, the amplitudes of the emitted waves, i.e., those occurring after the disintegration, are comparable with the amplitude of the initial wave, Second, the emission is associated with the transition from one type of discontinuity to the other, rater than with the oscillation of the discontinuity surface.

Thus, the scenario of time evolution of a trans-Alfvénic shock wave suggested by us is in agreement with the viewpoint according to which the shock cannot exist as a stationary configuration.

[1] N. E. Kotchine, Rendiconti del Circolo Matematico di Palermo **50**, 305 (1926).

[2] H. A. Bethe, Office of Scientific Research and Development, Rep. No. 445 (1942).

[3] G. Ya. Lyubarskii and R. V. Polovin, Zh. Éksp. Teor. Fiz. **35**, 1291 (1959) [Sov. Phys. JETP **8**, 901 (1959)].

[4] V. V. Gogosov, Prikl. Mat. Mekh. **25**, 108 (1961).

[5] L. D. Landau, Zh. Éksp. Teor. Fiz. **14**, 240 (1944).

[6] R. Courant and K. O. Fridrichs, Supersonic Flows and Shock Waves, Interscience Publ., New York (1948).

[7] P. Lax, Commun. Pure Appl. Math. **10**, 537 (1957).

[8] A. I. Akhiezer, G. Ya. Lyubarskii, and R. V. Polovin, Zh. Éksp. Teor. Fiz. **35**, 731 (1959) [Sov. Phys. JETP **8**, 507 (1959)].

[9] V. M. Kontorovich, Zh. Éksp. Teor. Fiz. **35**, 1216 (1959) [Sov. Phys. JETP **8**, 851 (1959)].

[10] S. I. Syrovatskii, Zh. Éksp. Teor. Fiz. **35**, 1466 (1959) [Sov. Phys. JETP **8**, 1024 (1959)].

[11] G. Ya. Lyubarskii and R. V. Polovin, Zh. Éksp. Teor. Fiz. **36**, 1272 (1959) [Sov. Phys. JETP **9**, 902 (1959)].

[12] R. V. Polovin and K. P. Cherkasova, Zh. Éksp. Teor. Fiz. **41**, 263 (1961) [Sov. Phys. JETP **14**, 190 (1961)].

[13] K. P. Cherkasova, Zh. Prikl. Mekh. Tekh. Fiz. No. 6, 169 (1961).

[14] A. Kantrovitz and H. Petschek, in Plasma Physics in Theory and Application, ed. by W. B. Kunkel, McGraw-Hill, New York (1966), p. 148.

[15] C. C. Wu and C. F. Kennel, Phys. Rev. Lett. **68**, 56 (1992).

[16] C. C. Wu and C. F. Kennel, Phys. Fluids B **5**, 2877 (1993).

[17] C. C. Wu, J. Geophys. Res. **93**, 987 (1988).

[18] L.-N. Hau and B. U. Ö. Sonnerup, J. Geophys. Res. **94**, 6539 (1989).

[19] C. F. Kennel, R. D. Blandford, and C. C. Wu, Phys. Fluids B **2**, 987 (1990).

[20] C. C. Wu, J. Geophys. Res. **95**, 8149 (1990).

[21] P. Germain, Rev. Mod. Phys. **32**, 951 (1960).

[22] A. G. Kulikovskii and G. A. Liubimov, Prikl. Mat. Mekh. **25**, 125 (1961).

[23] J. E. Anderson, Magnetohydrodynamic Shock Waves, M.I.T. Press, Cambridge, Massachusetts (1963).

[24] S. A. Markovskii and B. V. Somov, Space Sci. Rev. **78**, 443 (1996).

[25] Z. B. Roikhvarger and S. I. Syrovatskii, Zh. Éksp. Teor. Fiz. **66**, 1338 (1974) [Sov. Phys. JETP **39**, 654 (1974)].

[26] J. Bazer and W. B. Ericson, Astrophys. J. **129**, 758 (1959).

[27] S. P. D'yakov, Zh. Éksp. Teor. Fiz. **27**, 288, (1954).

[28] V. M. Kontorovich, Zh. Éksp. Teor. Fiz. **33**, 1525 (1958) [Sov. Phys. JETP **6**, 1179 (1958)].

[29] C. S. Gardner, Phys. Fluids **6**, 1366 (1963).

[30] N. M. Kuznetsov, Zh. Éksp. Teor. Fiz. **88**, 470, (1985) [Sov. Phys. JETP **61**, 275 (1985)].

[31] A. G. Kulikovskii and G. A. Lyubimov, Magnetohydrodynamics, Fizmatgiz, Moscow (1962).

[32] R. W. Griffits, R. J. Sandeman, and H. G. Hornung, J. Phys. D **8**, 1681 (1975).

[33] T. I. Mishin, A. P. Bedin, N. I. Yushenkova, G. E. Skvortsov, and A. P. Ryasin, Zh. Tekh. Fiz. **51**, 2315 (1981) [Sov. Phys. Tech. Phys. **26**, 1363 (1981)].

[34] L. D. Landau and E. M. Lifshitz, Fluid Dynamics, Nauka, Moscow (1986), p. 476.

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

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Filamentation and stimulated Brillouin scattering in a turbulent plasma

K. N. Ovchinnikov, V. P. Silin, and S. A. Uryupin

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

(Submitted 29 May 1997)
Zh. Éksp. Teor. Fiz. **113**, 629–645 (February 1998)

This paper presents a theory of filamentation and stimulated Brillouin scattering (SBS) of high-frequency electromagnetic radiation in a weakly collisional plasma with ion-acoustic turbulence. When the square of the wavelength of the plasma perturbations is less than the product of the two mean free path lengths of an electron with respect to its collisions with turbulent fluctuations and with electrons, the influence of cold highly collisional electrons on the parametric instabilities becomes apparent. It is shown that the plasma turbulence lowers the filamentation threshold, and the SBS threshold can be either lowered or raised. The dependence of the SBS and filamentation thresholds on the electron mean free path length in the turbulent plasma and on the anisotropy of the plasma turbulence is determined. The corresponding dependence of the spatial scale of the most efficiently growing filaments and their spatial growth rate are found. © *1998 American Institute of Physics.* [S1063-7761(98)01902-7]

1. INTRODUCTION

The effect on a plasma produced by high-intensity electromagnetic radiation with a frequency above the Langmuir frequency of the electrons to first order in the radiation intensity leads to a low-frequency modulation of the charge density. The density modulation arises both because of the effect of the ponderomotive force and because of heat transport and the heating of the electrons, which arise in the presence of inverse bremsstrahlung absorption of the radiation.[1,2] A change in the electron density in a high-intensity radiation field is the cause of such characteristic parametric instabilities of nonlinear phenomena as filamentation and SBS. Since the density perturbation associated with the collisional heating of the electrons depends strongly on the characteristics of the heat transport, for those conditions when the primary cause of the density modulation is the inverse bremsstrahlung the filamentation and SBS mechanisms depend on the characteristics of the heat transport (see, for example, Refs. 3, 4). The heat transport was initially considered in the collisional plasma limit.[5] However, it later became clear that for the conditions usually assumed to be collisionless, when the characteristic scale of the plasma perturbations is small compared with the mean free path of the electrons, there are always slow subthermal highly collisional electrons (the velocity of which is less than thermal). Taking this circumstance into account led to the development of the kinetic theory of a weakly collisional plasma, in which new filamentation and SBS mechanisms are discovered. Unusual filamentation mechanisms were revealed in Ref. 6 on the basis of numerical calculations in such a plasma and approximate scaling of the nonlocal electron transport was proposed. An analytical theory of a weakly collisional plasma, necessary for describing filamentation, was constructed in Ref. 7 and the exact scaling of the heat transport in a plasma with a Maxwellian electron distribution was determined. The corresponding SBS theory was constructed in Refs. 8, 9. It was shown in Ref. 9 that if the concept of nonlocal transport[6] is used in a weakly collisional plasma to describe parametric instabilities, then the heat transport that characterizes the damping of ionic sound differs qualitatively from the heat transport that determines the nonlinear interaction which forms the basis of filamentation theory. The theory of filamentation in a weakly collisional plasma in a fairly strong radiation field, when the heating radiation makes the electron distribution considerably different from Maxwellian, was developed in Refs. 10, 11. We us emphasize that the achievements of the theory of parametric instabilities of a weakly collisional plasma pointed out here were obtained through consideration of the properties of a laminar plasma.

On the other hand, the theory of nonlocal electron transport of heat in a weakly collisional plasma with ion-acoustic turbulence, when the heating of the plasma itself occurs because of the inverse bremsstrahlung absorption of the high-frequency field and when the plasma perturbation is caused by a low-frequency electric potential, was recently formulated (see Refs. 12, 13). The results of Refs. 12, 13 open up the possibility for constructing a theory of the parametric instabilities—filamentation and SBS—under conditions not previously considered when, first, the mean free path of the electrons l_t, which determines the isotropization of the electron distribution function, is not associated with electron–ion collisions but rather with their scattering at turbulent pulsations and, secondly, which is most significant, the absolute magnitude of the wave vector k of the plasma perturbations turns out to be so large that

$$Zll_tk^2 > 1. \tag{1.1}$$

Here Zl is the mean free path of a thermal electron with respect to its collisions with electrons. This inequality defines the conditions for which one must speak of a weakly collisional turbulent plasma. Since $l_t \ll l$, then in the presence of turbulence the plasma turns out to be weakly collisional in the case of perturbations whose spatial scale is much smaller than the corresponding scale of the perturbations of the laminar plasma. This corresponds to the fact that the condition in the collisionless situation for a turbulent plasma is

$$kl_t > 1. \tag{1.2}$$

The kinetic equation for the electrons in a plasma with ion-acoustic turbulence, located in a high-frequency electromagnetic field with frequency ω_0 greater than the electron Langmuir frequency ω_L, constitutes the starting point for the subsequent consideration. On the basis of the expression established earlier (see, for example, Ref. 14) for the collision integral of the electrons with anisotropic ion-acoustic flustuations of the charge density expressions are derived for the low-frequency perturbations of the electron density that arise both because of low-frequency oscillations of the electric potential and also because of the inverse bremsstrahlung. With the expression for the density modulation due to radiation absorption in the presence of electron–ion collisions taken into account the spatial gain factor of the filament is found. The spatial scale of the perturbations having the maximum gain factor is determined. The explicit dependence of the filamentation threshold on the plasma parameters is found. It is shown that in a turbulent plasma the spatial growth rate of the filament is a factor of $(l/l_t)^{2/9} > 1$ greater than in a laminar plasma, and the scale of the most rapidly growing perturbations and the filamentation threshold are smaller by factors of $(l_t/l)^{1/9} < 1$ and $(l_t/l)^{2/7} < 1$, respectively. Because of the anisotropy of the turbulent noise the threshold value of the radiation intensity, the spatial growth rate and the scale of the most rapidly growing perturbation depend on the angle between the wave vector of the perturbations and the direction of the density of the turbulence-producing force.

The effect of perturbations of the electron density, caused both by fluctuations of the electric potential and by heating of the electrons in electron–ion collisions, on the SBS threshold is investigated. The modulation of the electron density because of collisional absorption leads to a lowering of the instability threshold by a factor of $(l_t/l)^{2/7} < 1$. Conversely, the modulation of the density caused by fluctuations of the electric potential leads to an extremely large increase of the instability threshold by a factor of $(l/l_t)^{5/7} \gg 1$. This is caused by the anomalous increase of the damping of the ion-acoustic perturbations due to collisions of subthermal electrons. Since the regions in wavenumber space in which these two causes of density perturbation appear do not coincide, the effect of the turbulent noise on SBS can be either to raise or to lower the instability threshold. In both cases, however, the SBS threshold, like the filamentation threshold, has considerable anisotropy caused by the anisotropy of the nonlocal electron transport of heat in a plasma with anisotropic turbulent noise.

2. ELECTRON KINETICS IN A TURBULENT PLASMA

As is known, the quasistationary electric field $\nabla\phi$ and the electron pressure gradient ∇p are causes of the development of ion-acoustic turbulence in a nonisothermal plasma. If the effective force density acting on the electrons

$$\mathbf{R} = -en\nabla\phi - \nabla p, \tag{2.1}$$

greatly exceeds the value corresponding to the threshold for the excitation of the ion-acoustic instability

$$R \gg R_{\rm th} = \nu_{ei} n m v_s, \tag{2.2}$$

then the level of the low-frequency charge density fluctuations differs considerably from thermal. In Eqs. (2.1) and (2.2) e is the charge, m is the mass, $p = nmv_T^2$, v_T is the thermal velocity of the electrons, v_s is the velocity of sound, n is the electron density,

$$\nu_{ei} = \nu_{ei}(v_T) = \frac{4\sqrt{2\pi}\, Ze^4 n}{3m^2 v_T^3}\,\Lambda \tag{2.3}$$

is the collision frequency of the electrons with the ions having a charge $e_i = Z|e|$, and Λ is the Coulomb logarithm. Under the conditions of the inequality (2.2) for describing slow low-frequency electron motions, occurring during a time greater than the period of the Langmuir oscillations of the electrons, it is sufficient to take account of only their scattering by ion-acoustic fluctuations of the charge density and the usual electron–ion collisions can be ignored. If a quasistationary turbulent state is established because of the Cerenkov interaction of waves with electrons and ions and the stimulated scattering of waves by ions, the quasilinear collision integral that describes the effect of turbulence on the small deviation δf of the electron distribution function from Maxwellian f_m has the form[14]

$$\mathrm{St}_{QL}(\delta f) = \nu_t \frac{v_T^3}{v^3}\left\{\frac{\partial}{\partial\xi}\left[(1-\xi^2)\mathrm{X}(\sqrt{1-\xi^2})\,\frac{\partial}{\partial\xi}\,\delta f\right] + \frac{\Xi(\sqrt{1-\xi^2})}{1-\xi^2}\,\frac{\partial^2}{\partial\varphi^2}\,\delta f\right\}, \tag{2.4}$$

where $\xi = \cos\theta$, θ and φ are the angles of the velocity vector in the spherical coordinate system with the symmetry axis oriented along the direction of the effective force density $\mathbf{R}$. The form of the turbulent collision frequency ν_t and the functions X and Ξ depends on the absolute magnitude R of the force:

$$\nu_t = \sqrt{9\pi/8}\; R/nmv_s(1+\delta), \quad R \ll R_{NL}, \tag{2.5}$$

$$\nu_t = \sqrt{9\pi/8}\sqrt{RR_{NL}}/nmv_s, \quad R \gg R_{NL}, \tag{2.6}$$

where $R_{NL} = (1+\delta^2)nmv_s\omega_{Li}^2 r_D^2/6\pi\tau_{Di}^2$, ω_{Li} and r_{Di} are, respectively, the Langmuir frequency and Debye radius of the ions, r_D is the Debye radius of the electrons, δ is the ratio of the ion and electron Cerenkov damping rates of sound. The dimensionless anisotropic collision frequencies X and Ξ, realized at large and small R, are shown in Figs. 1 and 2.

Let us consider the interaction of a completely ionized turbulent plasma with a high-frequency electromagnetic field of the form

$$\frac{1}{2}\,\mathbf{E}(\mathbf{r},t)\exp(-i\omega_0 t) + \text{c.c.}, \tag{2.7}$$

where the amplitude $\mathbf{E}(\mathbf{r},t)$ varies slightly during the period $2\pi/\omega_0$. We assume that the frequency ω_0 exceeds both the Langmuir frequency of the electrons ω_L and the effective

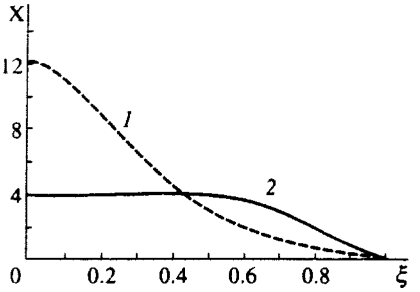


FIG. 1. Dimensionless effective collision frequency determining the relaxation of the electron velocity vector with respect to the polar angle. Curve *1* corresponds to $R/R_{NL}=0.03$ and curve *2* is valid for $R\gg R_{NL}$.

collision frequency of electrons with ions and turbulent noise. Under these conditions the effect of low-frequency long-wave fluctuations of the charge density on the vibrational motion of the electrons with frequency ω_0 is strongly suppressed.[15,16] Conversely, the comparatively rare electron–ion collisions with small impact parameters of order $v_T/\omega_0\ll r_D$ produce the primary influence on the rapidly varying motion of the electrons. The effect of these collisions on the deviation of the distribution function from Maxwellian can be described by an electron–ion collision integral of the form

$$\mathrm{St}(\delta f)=\sqrt{\frac{9\pi}{8}}\,\nu_{ei}(v)\,\frac{\partial}{\partial v_i}\,(v^2\delta_{ij}-v_iv_j)\,\frac{\partial}{\partial v_j}\,\delta f, \tag{2.8}$$

where the frequency $\nu_{ei}(v)$ depends on the velocity according to Eq. (2.3), in which v_T is replaced by v and the maximum impact parameter of the Coulomb logarithm is equal to v_T/ω_0.

We will restrict consideration to comparatively low radiation intensities when the amplitude of the electron quiver velocity $v_E=|e\mathbf{E}/m\omega_0|$ in the electro-magnetic field is less than $v_T/\sqrt{Z}\ll v_T$. Then the effect of the high-frequency field on the electron distribution function can be taken into account using perturbation theory. Moreover, the effect of the radiation ponderomotive force on the electrons will be assumed to be small compared with the effect of the effective force density $\mathbf{R}$ given by (2.1). This last assumption makes it possible to ignore the effect of the radiation on turbulence. Under these conditions we have an equation of the form

$$\begin{aligned}&\frac{\partial}{\partial t}\,\delta f+\mathbf{v}\cdot\frac{\partial}{\partial\mathbf{r}}\,\delta f-\mathrm{St}(\delta f,f_m)-\mathrm{St}(f_m,\delta f)-\mathrm{St}_{QL}(\delta f)\\&=\frac{e}{m}\,(\nabla\delta\phi)\cdot\frac{\partial f_m}{\partial\mathbf{v}}+\frac14\left(\frac{\partial}{\partial\mathbf{r}}\,v_E^2\right)\cdot\frac{\partial f_m}{\partial\mathbf{v}}+\frac18\,\frac{\partial^2 f_m}{\partial v_i\partial v_j}\\&\times\left(\frac{\partial}{\partial t}+\mathbf{v}\cdot\frac{\partial}{\partial\mathbf{r}}\right)V_{ij}-\sqrt{\frac{\pi}{8}}\,v_E^2v_T\nu_{ei}\,\frac{\partial}{\partial v_i}\left(f_m\,\frac{v_i}{v^3}\right)\\&-\frac32\,\sqrt{\frac{\pi}{8}}\left(V_{ij}-\frac23\,\delta_{ij}v_E^2\right)v_T\nu_{ei}\,\frac{\partial}{\partial v_i}\left(f_m\,\frac{v_j}{v^3}\right)\end{aligned} \tag{2.9}$$

for the nonequilibrium correction δf to the distribution function, caused by the perturbation $\delta\phi$ of the potential of the quasistationary electric field and by the high-frequency radiation, in a linear approximation in terms of the radiation intensity. Here V_{ij} is the quiver velocity tensor,

$$V_{ij}=v_{Ei}v_{Ej}^{*}+v_{Ej}v_{Ei}^{*}, \tag{2.10}$$

and the electron–electron collision integral $\mathrm{St}(\delta f,f_m)+\mathrm{St}(f_m,\delta f)$ is taken into account. Equation (2.9) makes it possible to investigate the slow low-frequency motions of electrons in a turbulent plasma interacting with low-frequency radiation.

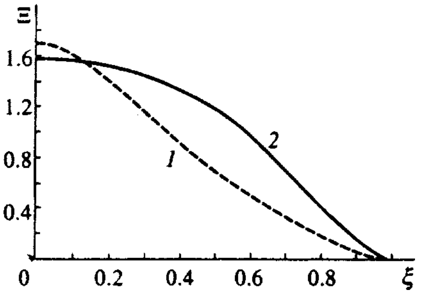


FIG. 2. Dimensionless effective collision frequency determining the relaxation of the electron velocity vector with respect to the azimuth angle. Curves *1* and *2* correspond to the same parameters as in Fig. 1.

3. ELECTRON DENSITY PERTURBATION

Assuming that the quiver velocity tensor V_{ij} and the perturbation of the potential $\delta\phi$ are low-frequency long-wave perturbations of the form

$$\sim\exp(i\mathbf{kr}-i\omega t),$$

we use Eq. (2.9) to describe the perturbation of the electron density δn with frequency ω and wave vector $\mathbf{k}$. Introducing the function δf_c, which is related to δf by the expression

$$\delta f=\delta f_c-\frac{e\delta\phi}{mv_T^2}\,f_m-If_m+\frac18\,V_{ij}\,\frac{\partial^2 f_m}{\partial v_i\partial v_j}, \tag{3.1}$$

where $I=v_E^2/4v_T^2$, we have

$$\begin{aligned}&-i(\omega-\mathbf{k}\cdot\mathbf{v})\delta f_c-\mathrm{St}(\delta f_c,f_m)-\mathrm{St}(f_m,\delta f_c)-\mathrm{St}_{QL}(\delta f_c)\\&=-\sqrt{\frac{\pi}{8}}\,v_E^2v_T\nu_{ei}\,\frac{\partial}{\partial v_i}\left(f_m\,\frac{v_i}{v^3}\right)-\frac32\,\sqrt{\frac{\pi}{8}}\\&\times\left(V_{ij}-\frac23\,\delta_{ij}v_E^2\right)v_T\nu_{ei}\,\frac{\partial}{\partial v_i}\left(f_m\,\frac{v_j}{v^3}\right)-i\omega If_m\\&-i\omega\,\frac{e\delta\phi}{mv_T^2}\,f_m+\mathrm{St}_{QL}\left(\frac18\,V_{ij}\,\frac{\partial^2 f_m}{\partial v_i\partial v_j}\right),\end{aligned} \tag{3.2}$$

for finding δf_c. Since Eq. (3.2) is linear, all of the inhomogeneous terms on the right side contribute independently to δf_c. In particular, for $\omega<kv_T$ and $kv_T>\nu_t$ the term containing $\delta\phi$ gives a contribution to δf_c that leads to collisionless Landau damping. Identifying only the collisionless contribution from δf_c and taking Eq. (3.1) into consideration, we write the following expression for the density perturbation:

$$\delta n_T = -n\left[I + \frac{e\delta\phi}{mv_T^2}\left(1 + i\,\frac{\omega}{kv_T}\,\sqrt{\frac{\pi}{2}}\right)\right], \tag{3.3}$$

which is caused by the effect of the potential perturbation $\delta\phi$ and the ponderomotive force on the bulk of electrons with velocities $v \sim v_T$. In this case the corrections to the effect of the ponderomotive force, caused by the term $i\omega I f_m$, are ignored since they are small for the conditions to be discussed below.

Let us now consider Eq. (3.2) for subthermal collisional electrons for the conditions

$$\omega \ll \nu_t(v) = \nu_t v_T^3/v^3, \quad kv \ll \nu_t(v)$$

and ignoring the term $i\omega I f_m$. We look for a solution in the form $\delta f_c = \delta f_0 + \delta f_a$, where $\delta f_0 = \int d\Omega\,\delta f_c/4\pi$, $d\Omega = \sin\theta d\theta d\varphi$, and the small anisotropic addition is found from the equation

$$i\mathbf{k}\cdot\mathbf{v}\delta f_0 = \mathrm{St}_{QL}(\delta f_a). \tag{3.4}$$

Averaging the original Eq. (3.2) over the angles of the velocity vector and using the solution of Eq. (3.4), we find that δf_0 approximately satisfies

$$\begin{aligned} -i\omega\delta f_0 &+ \frac{v^2}{4\nu_t(v)}\,(\beta_\parallel k_z^2 + \beta_\perp k_\perp^2)\,\delta f_0 \\ &- \frac{1}{Z}\,\nu_{ei}\,\frac{v_T^2}{v^2}\,\frac{\partial}{\partial v}\left(v^2\,\frac{\partial}{\partial v}\,\delta f_0\right) = -\,\frac{i\omega}{mv_T^2}\,e\delta\phi_{fm}(v) \\ &- \sqrt{\frac{\pi}{8}}\,\frac{v_E^2}{vv_T}\,\nu_{ei} f_m(v)\,\delta\!\left(\frac{v^2}{2v_T^2}\right), \end{aligned} \tag{3.5}$$

where $k_z = (\mathbf{k}\cdot\mathbf{n})$, $k^2 = k_z^2 + k_\perp^2$, $\mathbf{n} = \mathbf{R}/R$ is the unit vector of the effective force density, and the coefficients $\beta_\parallel$ and $\beta_\perp$ are equal to

$$\beta_\parallel = 0.18, \quad \beta_\perp = 0.02, \quad R \ll R_{NL}, \tag{3.6}$$

$$\beta_\parallel = 0.25, \quad \beta_\perp = 0.80, \quad R \gg R_{NL}. \tag{3.7}$$

Equation (3.5) is analyzed below under the assumption that the parameter

$$N_t = Z\,\frac{v_T^2}{\sqrt{2}\,\nu_t\nu_{ei}}\,(\beta_\parallel k_z^2 + \beta_\perp k_\perp^2) \tag{3.8}$$

is much greater than unity. In this limit the contribution from δf_0 to the density perturbation is caused by subthermal electrons with velocities $v \sim v_T N_t^{-1/7} < v_T$. With this fact taken into account two assumptions are made in the writing of Eq. (3.5). First, the electron–electron collision integral is written for slow electrons with $v < v_T$ and, secondly, on the right side of Eq. (3.5) the 1 is ignored compared with the δ function. Following Refs. 7, 13, in the limit $N_t \gg 1$ and for

$$\omega \ll \nu_t(v)[kv/\nu_t(v)]^2, \tag{3.9}$$

one can write the exact solution of Eq. (3.5) asymptotically. Since the density perturbation is determined by particles with velocities $v \sim v_T N_t^{-1/7} < v_T$, and $N_t \gg 1$, the results presented below hold under the following restriction on the frequency (see Eq. (3.9)) and the wave vector of the perturbations:

$$\omega \ll N_t^{2/7}\nu_{ei}/Z, \tag{3.10}$$

$$kl_t(\theta_k) \gg \sqrt{l_t(\theta_k)/Zl}. \tag{3.11}$$

The notations

$$l_t(\theta_k) = l_t\beta(\theta_k)/\sqrt{2}, \tag{3.12}$$

$$l_t = v_T/\nu_t \ll l = v_T/\nu_{ei}, \tag{3.13}$$

$$\beta(\theta_k) = \beta_\parallel\cos^2\theta_k + \beta_\perp\sin^2\theta_k, \tag{3.14}$$

are used here, where θ_k is the angle between the wave vector of the perturbations and the direction of the effective force density causing the turbulence. With these replacements taken into account, integrating the solution of Eq. (3.5) with respect to the magnitude of the velocity and adding the contribution from the thermal electrons δn_T (3.3), we find $\delta n = \delta n_I + \delta n_\phi$,

$$\delta n_I = -In(1 + \beta_0 Z N_t^{-2/7}), \tag{3.15}$$

$$\delta n_\phi = -\,\frac{e\delta\phi}{mv_T^2}\,n\left(1 + i\sqrt{\frac{\pi}{2}}\,\frac{\omega}{kv_T} + i\beta_{1/2}\,\frac{\omega}{\nu_{ei}}\,ZN_t^{-5/7}\right), \tag{3.16}$$

where $\beta_0 = 1.16$, $\beta_{1/2} = 0.82$. The relations (3.15) and (3.16) serve as the starting point for the further consideration of such electromagnetic phenomena as filamentation and stimulated Brillouin scattering.

In Eqs. (3.15) and (3.16) the weakly collisional contribution of the subthermal electrons is found to be greater than the collisionless for not too large values of k. At the same time the condition $N_t \gg 1$ requires fairly large values of k. This leads to the following two relations. First, the effect of the ponderomotive force in Eq. (3.15) is small if

$$Z^{7/2} > k^2 Zll_t(\theta_k) \gg 1. \tag{3.17}$$

Secondly, the weakly collisional contribution to Eq. (3.16) exceeds the collisionless Landau damping if

$$[Zl/l_t(\theta_k)]^{7/3} > k^2 Zll_t(\theta_k) \gg 1. \tag{3.18}$$

Here the condition $Z \gg 1$ is necessary to satisfy Eq. (3.17). In contrast to this no such condition is required in the case of Eq. (3.18) since $Z \gg 1_t(\theta_k)/1$ always holds.

On the other hand, in a turbulent plasma it can happen that the perturbation of the electron density determined by the radiation intensity, is related to the ponderomotive force, and the perturbation (3.16) is related to the subthermal collisional electrons. This is true in the case when the condition

$$[Zl/l_t(\theta_k)]^{7/3} > k^2 Zll_t(\theta_k) > Z^{7/2} \tag{3.19}$$

is satisfied for the wave vector $\mathbf{k}$. This last condition is possible for an ionization multiplicity that is not too large,

$$Z < [l/l_t(\theta_k)]^2. \tag{3.20}$$

In concluding this section we must point out that in writing Eq. (3.5) we completely ignored the time dependence of v_T. If it is assumed that the heating radiation is sufficiently uniform and heat loss can be ignored, then the rate of variation of the thermal velocity with time is characterized by the equation

$$\frac{d}{dt}\,v_T^2=\frac{1}{3}\,v_{E_0}^2\nu_{ei}(v_T); \tag{3.21}$$

here $v_{E_0}=|e\mathbf{E}_0/m\omega_0|$, where $\mathbf{E}_0$ is the intensity of the spatially uniform radiation heating the plasma. Similarly to the condition (3.9), the time dependence of v_T in Eq. (3.5) can be ignored for the condition

$$\nu_t(v)[kv/\nu_t(v)]^2 \gg v_{E_0}^2\nu_{ei}(v_T)/v_T^2. \tag{3.22}$$

Taking into consideration that $v\sim v_T N_t^{-1/7}$, we can write the final relation in the form

$$v_T^2 N_t^{2/7}/Z \gg v_{E_0}^2. \tag{3.23}$$

Since $N_t \gg 1$, the condition (3.23) is weaker than the previously adopted restriction

$$v_{E_0}^2 \ll v_T^2/Z, \tag{3.24}$$

which allows us to ignore how electron heating due to inverse bremsstrahlung alters the electron distribution function.

4. RADIATION FILAMENTATION

In order to consider the filamentation phenomenon of laser radiation we assume in Eq. (3.15) that the electric field has the form

$$\mathbf{E}(\mathbf{r},t)=[E_0+\delta E_1\,\exp(iky)+\delta E_{-1}\times\exp(-iky)]\mathbf{e}_x\,\exp(ik_0 z), \tag{4.1}$$

where E_0 is the field intensity of the fundamental wave, $\delta E_{\pm 1}$ are the small amplitudes of the filamentation perturbations, k_0 is the wave number, and $\mathbf{e}_x$ is the unit radiation polarization vector. Then to first order we have

$$|\mathbf{E}|^2=|E_0|^2+(\delta E_1 E_0^* + E_0\delta E_{-1}^*)\exp(iky)+\text{c.c.} \tag{4.2}$$

In accordance with the previously obtained Eq. (3.15) a perturbation of the radiation field intensity leads to a perturbation of the electron density

$$\delta n=-\frac{n}{16\pi n_c m v_T^2}\,(E_0^*\delta E_1+E_0\delta E_{-1}^*)\times(1+\beta_0 Z N_t^{-2/7}(\theta_k)), \tag{4.3}$$

where $n_c=m\omega_0^2/4\pi e^2$ is the critical electron density and θ_k is the angle between the component of the wave number k and the anisotropy direction of the turbulent noise. The term in Eq. (4.3) that is independent of β_0 describes the modulation of the electron density caused by the action of the ponderomotive force. Conversely, the term containing β_0 characterizes the density modulation due to heating of the slow electrons in the presence of inverse bremsstrahlung, being a superposition of the fields of the test wave and the perturbations. In accordance with Maxwell's equations the amplitudes of the small electromagnetic field perturbations satisfy truncated equations of the form

$$\left(2ik_0\,\frac{d}{dz}-k^2\right)\delta E_1(z)=\frac{\omega_0^2}{c^2}\,\frac{\delta n}{n_c}\,E_0, \tag{4.4}$$

$$\left(-2ik_0\,\frac{d}{dz}-k^2\right)\delta E_{-1}^*(z)=\frac{\omega_0^2}{c^2}\,\frac{\delta n}{n_c}\,E_0^*. \tag{4.5}$$

We look for the solution of these equations in the form of a function increasing exponentially in the direction of the fundamental wave $\delta E_1\sim\delta E_{-1}^*\sim\exp(Gz)$. Taking Eq. (4.3) into account, we find

$$G^2(k)=\left(\frac{k}{k_0}\right)^2\left[q^2(k)-\frac{k^2}{4}\right], \tag{4.6}$$

$$q^2(k)=\frac{n\omega_0^2}{32\pi n_c^2 m v_T^2 c^2}\,|E_0|^2(1+\beta_0 Z N_t^{-2/7}) \tag{4.7}$$

from the system (4.4), (4.5) for the spatial gain factor of the filament. The effect of turbulence on filamentation appears for the conditions

$$\beta_0 Z \gg N_t^{2/7}\equiv N_0^{2/7}(k/k_0)^{4/7}\gg 1, \tag{4.8}$$

when the modulation of the electron density is caused by a heating of the slow electrons in the presence of inverse bremsstrahlung. Within the wave number interval

$$(\beta_0 Z)^{7/4}k_0/\sqrt{N_0}\gg k\gg k_0/\sqrt{N_0}\equiv[Zll_t(\theta_k)]^{-1/2}, \tag{4.9}$$

the spatial gain factor of the filament reaches the maximum value

$$G_m(\theta_k)=\frac{3k_0}{2\sqrt{5}}\,\frac{Z^{5/9}}{[k_0^2 ll_t(\theta_k)]^{2/9}}\left[\frac{5\beta_0}{112\pi}\,\frac{n\omega_0^2|E_0|^2}{n_c^2 m v_T^2 k_0^2 c^2}\right]^{7/9} \tag{4.10}$$

for the wave number

$$k_m(\theta_k)=\frac{k_0 Z^{5/18}}{[k_0^2 ll_t(\theta_k)]^{1/9}}\left[\frac{5\beta_0}{112\pi}\,\frac{n\omega_0^2|E_0|^2}{n_c^2 m v_T^2 k_0^2 c^2}\right]^{7/18} \tag{4.11}$$

which satisfies the inequalities $k_0\gg k_m\gg G_m$. The wave number k_m corresponding to the maximum spatial gain lies within the interval of wave numbers (4.9) being considered if the radiation and plasma parameters satisfy the conditions

$$(\beta_0 Z)^{9/2}\gg\frac{5\beta_0}{28}\,\frac{\omega_L^2}{c^2}\,\frac{v_{E_0}^2}{v_T^2}\,Z^2 ll_t(\theta_k)\gg 1. \tag{4.12}$$

For a plasma with multiply charged ions satisfying $Z\gg 1$ the conditions (4.12) hold over a wide interval of frequencies and radiation intensities. Let us assume that the characteristic scale of variation of the plasma density is equal to L. Then, using the estimate $G_m L=2\pi$, we find

$$|E_0|^2_{\text{th},\theta_k}=\frac{112\pi}{5\beta_0}\left(\frac{4\pi\sqrt{5}}{3k_0 L}\right)^{9/7}Z^{-5/7}[k_0^2 ll_t(\theta_k)]^{2/7}\,\frac{n_c^2}{n}\,m\left(\frac{k_0 c}{\omega_0}\,v_T\right)^2 \tag{4.13}$$

for the filamentation threshold. An estimation of the wave number at filamentation threshold, in turn, gives[7]

$$(k_m)_{\text{th}}=k_0\left(\frac{4\pi\sqrt{5}}{3k_0 L}\right)^{1/2}\simeq 1.22k_0\left(\frac{2\pi}{k_0 L}\right)^{1/2}. \tag{4.14}$$

Substituting this expression into the inequality (3.17), we obtain

$$Z^{7/2} \gg 60 Z l l_t(\theta_k)/\lambda_0 L \gg 1, \tag{4.15}$$

where $\lambda_0 = 2\pi/k_0$ is the pump wavelength. Let us compare Eqs. (4.10)–(4.13) with their analogs, obtained earlier in the theory of filamentation for a laminar plasma.[7,10,11] Filamentation in a plasma with ion-acoustic turbulence differs in two important ways from filamentation in a laminar plasma. First, the spatial gain factor of the filament in a turbulent plasma according to Eq. (4.10) is a factor of $(l/l_t)^{2/9}$ greater, the characteristic scale of the most efficiently increasing perturbations $\sim 1/k_m$ according to Eq. (4.11) is a factor of $(l_t/l)^{1/9}$ less, and the filamentation threshold according to Eq. (4.13) is a factor of $(l_t/l)^{2/7}$ lower because of the change in the electron mean free path. Secondly, the quantities $G_m(\theta_k)$, $k_m(\theta_k)$ and $|E_0|^2_{\mathrm{th},\theta_k}$ depend on the angle θ_k between the wave vector of the perturbations and the anisotropy direction of the turbulent noise because of the anisotropy of the electron scattering in a turbulent plasma. The appearance of anisotropy in the filamentation phenomenon depends on the level of turbulent noise. For example, for a comparatively low turbulence level with $R \ll R_{NL}$ (2.5), we have

$$\frac{G_m(\theta_k=0)}{G_m(\theta_k=\pi/2)} = 0.61, \quad \frac{k_m(\theta_k=0)}{k_m(\theta_k=\pi/2)} = 0.78,$$
$$\frac{|E_0|^2_{\mathrm{th},\theta_k=0}}{|E_0|^2_{\mathrm{th},\theta_k=\pi/2}} = 1.87. \tag{4.16}$$

Conversely, for $R \gg R_{NL}$ (2.6) the ratio of the quantities being considered for $\theta_k=0$ and $\theta_k=\pi/2$ is changed:

$$\frac{G_m(\theta_k=0)}{G_m(\theta_k=\pi/2)} = 1.29, \quad \frac{k_m(\theta_k=0)}{k_m(\theta_k=\pi/2)} = 1.14,$$
$$\frac{|E_0|^2_{\mathrm{th},\theta_k=0}}{|E_0|^2_{\mathrm{th},\theta_k=\pi/2}} = 0.72. \tag{4.17}$$

The necessary conditions for these filamentation mechanisms to occur are the inequalities (4.12), which are realized comparatively simply for a turbulent plasma with multiply charged ions, in which the laser radiation heating the target is absorbed efficiently.

5. STIMULATED BRILLOUIN SCATTERING

Let us examine the scattering of the test electromagnetic wave

$$\frac{1}{2}\mathbf{E}_0 \exp(-i\omega_0 t + i\mathbf{k}_0\mathbf{r}) + \text{c.c.} \tag{5.1}$$

with frequency $\omega_0 = (\omega_L^2 + k_0^2 c^2)^{1/2}$ by ion-acoustic oscillations of the electron density

$$\frac{1}{2}\delta n \exp(-i\omega t + i\mathbf{k}\mathbf{r}) + \text{c.c.} \tag{5.2}$$

with frequency $\omega \ll \omega_0$. As is usual, we represent the field of the scattered wave in the form

$$\frac{1}{2}\mathbf{E}_{-1} \exp(-i\omega_{-1} t + i\mathbf{k}_{-1}\mathbf{r}) + \text{c.c.}, \tag{5.3}$$

where the frequency $\omega_{-1} = (\omega_L^2 + k_{-1}^2 c^2)^{1/2}$ and the wave vector $\mathbf{k}_{-1}$ are related to ω_0, $\mathbf{k}_0$ and ω, $\mathbf{k}$ by the expressions

$$\omega_{-1} = \omega_0 - \omega, \quad \mathbf{k}_{-1} = \mathbf{k}_0 - \mathbf{k}. \tag{5.4}$$

Assuming the field of the fundamental wave $\mathbf{E}_0$ is given, we write a system of two linear equations of the form

$$\left[\omega_{-1}^2 - k_{-1}^2 c^2 - \omega_L^2\left(1 + \frac{i\nu_{ei}}{\omega_{-1}}\right)\right]\mathbf{E}_{-1}^* = \frac{\delta n}{2n}\,\omega_L^2 \mathbf{E}_0^*, \tag{5.5}$$

$$(\omega^2 - \omega_s^2 + 2i\omega\gamma_s)\delta n = \frac{e^2 n}{4m^2 v_T^2}\,\omega_s^2\left(\frac{1}{\omega_0^2} + \frac{1}{\omega_{-1}^2}\right) \times (1 + \beta_0 Z N_t^{-2/7})\mathbf{E}_0\mathbf{E}_{-1}^*, \tag{5.6}$$

for the amplitude of the scattered wave field and for the small amplitude of the perturbation of the electron density, where $\omega_s = k v_s/\sqrt{1 + k^2 r_D^2}$ is the frequency and γ_s is the sound damping rate. The expression for the perturbation of the electron density (3.15), arising both because of the effect of the ponderomotive force and also because of the inverse bremsstrahlung absorption of the field, was used for writing the right side of Eq. (5.6). In accordance with Sec. 3 above, this expression holds under the condition $N_t = Zk^2 l l_t(\theta_k) \gg 1$. The sound damping rate entering into Eq. (5.6) consists of two parts $\gamma_s = \gamma_i + \gamma_e$, caused by the ion damping γ_i and electron damping γ_e. Ion damping arises as a consequence of the Cerenkov interaction of sound with the ions and also because of ion–ion collisions:

$$\gamma_i = \sqrt{\frac{\pi}{8}}\,\frac{\omega_s^4}{k^3 v_{T_i}^3}\exp\left(-\frac{\omega_s^2}{2k^2 v_{T_i}^2}\right) + \gamma_{ii}, \tag{5.7}$$

where v_{T_i} is the thermal velocity of the ions, and the contribution to the damping rate from ion–ion collisions depends on the relationship between the ion–ion collision frequency ν_{ii} and the sound frequency

$$\gamma_{ii} = 0.64 k^2 v_{T_i}^2/\nu_{ii}, \quad \nu_{ii} \gg \omega_s,$$
$$\gamma_{ii} = 0.8\nu_{ii} k^2 v_{T_i}^2/\omega_s^2, \quad \omega_s \gg \nu_{ii}, \tag{5.8}$$

$$\nu_{ii} = \frac{4\sqrt{\pi} e^4 n}{3 m_i^2 v_{T_i}^3}\, Z^3 \Lambda_i, \tag{5.9}$$

m_i is the ion mass, and Λ_i is the ion Coulomb logarithm. The effect of ion-acoustic turbulence on the motion of the thermal ions is assumed to be small; this is valid for $\nu_{ii} \gg \nu_t \omega_{Li}^2/\omega_L^2$. Finally, using the expression for the time-dependent perturbation of the electron density (3.16), we write the sound damping rate by electrons[13]:

$$\gamma_e = \omega_s \frac{\omega_{Li}}{\omega_L}\left(\frac{\omega_s}{k v_s}\right)^3\left[\sqrt{\frac{\pi}{8}} + \frac{\beta_{1/2}}{2}(Zkl)^{2/7}(k l_t(\theta_k))^{-5/7}\right], \tag{5.10}$$

where the term $\sqrt{\pi/8}$ refers to taking the collisionless Landau damping into account.

Introducing the notations

$$\Delta=\frac{c^2}{2\omega_0}(k^2-2\mathbf{k}\cdot\mathbf{k}_0),\quad \gamma_E=\frac{\nu_{ei}\omega_L^2}{2\omega_0^2},\tag{5.11}$$

and ignoring the small difference between the frequencies ω_{-1} and ω_0, we obtain from the system of Eqs. (5.5) and (5.6) the dispersion equation for analyzing stimulated Brillouin scattering in a plasma with ion-acoustic turbulence:

$$-(\omega+\Delta+i\gamma_E)(\omega^2-\omega_s^2+2i\omega\gamma_s)=\frac{\omega_s^2}{8\omega_0 r_D^2}v_{E_0}^2$$
$$\times\{1+\beta_0 Z^{5/7}[k^2 l l_t(\theta_k)]^{-2/7}\}.\tag{5.12}$$

Here $\omega=\omega_s+\delta\omega$, and the frequency shift $\delta\omega$ is relatively small, $\delta\omega\ll\omega_s$. Moreover, since we have $|\mathbf{k}_{-1}|\simeq|\mathbf{k}_0|$, then $|\Delta|\ll\omega_0$ holds. In order to investigate the SBS we assume that the frequency ω has a small imaginary part $\gamma=\mathrm{Im}\,\omega$. The growth rate γ is maximum when the frequency shift satisfies $\delta\omega=0$, and the shift satisfies $\Delta=-\omega_s$. Under these conditions we find

$$\gamma=-\frac{1}{2}(\gamma_E+\gamma_s)+\left\{\frac{1}{4}(\gamma_E-\gamma_s)^2+\frac{\omega_s v_{E_0}^2}{16\omega_0 r_D^2}\right.$$
$$\left.\times[1+\beta_0 Z^{5/7}(k^2 l l_t(\theta_k))^{-2/7}]\right\}^{1/2}\tag{5.13}$$

from Eq. (5.12). According to Eq. (5.13) absolute instability of SBS in a turbulent plasma arises when the square of the electron quiver velocity in the electromagnetic field is greater than

$$(v_{E_0}^2)_{\mathrm{th}}=8\nu_{ei}\frac{v_T^2}{\omega_0\omega_s}[1+\beta_0 Z^{5/7}(k^2 l l_t(\theta_k))^{-2/7}]^{-1}$$
$$\times\left\{\gamma_i+\gamma_L\left[1+\sqrt{\frac{2}{\pi}}\beta_{1/2}(Zkl)^{2/7}(kl_t(\theta_k))^{-5/7}\right]\right\},\tag{5.14}$$

where $\gamma_L=\sqrt{\pi/8}(\omega_{L_i}/\omega_L)\omega_s^4(kv_s)^{-3}$, and the growth rate γ_i is described by Eqs. (5.7)–(5.9). We recall that an expression of the form (5.14) holds in the limit $N_t\gg1$ or $Zk^2 l l_t(\theta_k)\gg1$. The turbulence can affect the SBS threshold for two reasons. First, because of the effect of turbulence on the density perturbation because of the inverse bremsstrahlung. This effect is evident in the interval of wave numbers satisfying the condition (3.17). Since $l\gg l_t(\theta_k)$, for such wave numbers the threshold for the appearance of SBS is a factor of $[l_t(\theta_k)/l]^{2/7}$ lower than in a laminar plasma.[9] Because of the anisotropy of the electron scattering by turbulent noise the threshold value $|E_0|^2_{\mathrm{th},\theta_k}$ depends on the angle between the scattering wave vector and the direction of the effective force density producing the turbulence. The degree of anisotropy of the SBS threshold is characterized by Eqs. (4.16) and (4.17), the same as for the filamentation threshold.

Secondly, turbulence affects the SBS threshold because of a change in the electron damping of sound. This effect can appear within the wave number interval (see Eq. (3.18))

$$\left(\frac{2}{\pi}\beta_{1/2}^2\right)^{7/6}Z^{2/3}[l/l_t(\theta_k)]^{5/3}\gg kl\gg\sqrt{1/Zl_t(\theta_k)},\tag{5.15}$$

when the electron Landau damping is comparatively small. In order for the turbulent noise to alter the SBS significantly the condition ensuring smallness of the ion damping of sound,

$$\left(\frac{2}{\pi}\beta_{1/2}^2\right)^{7/6}Z^{2/3}[l/l_t(\theta_k)]^{5/3}(\gamma_L/\gamma_i)^{7/3}\gg kl,\tag{5.16}$$

must be satisfied in addition to the inequalities (5.15). If $\gamma_L>\gamma_i$ holds, then the restriction (5.16) is weaker than (5.15). In the wave number interval (5.15), (5.16), but outside the interval (3.17), the increase of the SBS threshold in a turbulent plasma caused by the effect of turbulence alone on sound damping compared with its value in a laminar plasma is characterized by the parameter $[l/l_t(\theta_k)]^{5/7}\gg1$. For these same conditions the degree of anisotropy of the SBS threshold is characterized by the expressions

$$|E_0|^2_{\mathrm{th},\theta_k=0}/|E_0|^2_{\mathrm{th},\theta_k=\pi/2}=0.21,\quad R\ll R_{NL},$$
$$|E_0|^2_{\mathrm{th},\theta_k=0}/|E_0|^2_{\mathrm{th},\theta_k=\pi/2}=2.3,\quad R\gg R_{NL}.\tag{5.17}$$

If conditions (3.17), (5.15) and (5.16) are satisfied simultaneously, then changes in either the effective ponderomotive force or the damping of sound in the plasma with turbulent noise affect the SBS threshold. The combined influence of these effects implies that the SBS threshold in the turbulent plasma is higher than in the laminar by the factor $[l/l_t(\theta_k)]^{3/7}\gg1$. In this case the degree of anisotropy of the SBS threshold is described by the relations

$$|E_0|^2_{\mathrm{th},\theta_k=0}/|E_0|^2_{\mathrm{th},\theta_k=\pi/2}=0.39,\quad R\ll R_{NL},$$
$$|E_0|^2_{\mathrm{th},\theta_k=0}/|E_0|^2_{\mathrm{th},\theta_k=\pi/2}=1.6,\quad R\gg R_{NL}.\tag{5.18}$$

The spatial growth rate $G=\mathrm{Im}\,k$ of the scattered wave realized under SBS conditions is of interest for experimental conditions. Let us consider the projection $G_{-1}=G\mathbf{k}_{-1}/k_{-1}\simeq G\mathbf{k}_{-1}/k_0$ of the spatial growth rate onto the direction of the scattered wave. The quantity G_{-1} has a maximum if the displacement of the real part of the frequency is equal to zero, $\delta\omega=0$. When $\delta\omega=0$ holds, just as for a laminar plasma, the spatial growth rate G_{-1} is described by[9]

$$G_{-1}=\frac{\nu_{ei}\omega_L^2}{2\omega_0 k_0 c^2}\left[\frac{v_{E_0}^2}{(v_{E_0}^2)_{\mathrm{th}}}-1\right],\tag{5.19}$$

where $(v_{E_0}^2)_{\mathrm{th}}$, however, depends on the turbulent noise level in accordance with Eq. (5.14).

6. CONCLUSION

The theory presented above for the simplest nonlinear phenomena arising during the interaction of high-intensity high-frequency electromagnetic radiation with a turbulent plasma has made it possible to identify new features of filamentation and SBS under those conditions of a weakly collisional plasma when the nonlocal nature of electron heat transport is important. We will dwell in greater detail on a discussion of filamentation and SBS mechanisms established above. Let us first discuss the conditions for the appearance of new filamentation mechanisms. According to Sec. 4 the wavenumber corresponding to the maximum spatial gain of the filament lies in the region where nonlocal theory is applicable if the radiation intensity satisfies the inequality (4.12). On the other hand, in order to ignore the evolution of the electron distribution with time the radiation intensity must be less than the value defined by the inequality (3.24). The simultaneous fulfilment of these inequalities is possible if

$$N_p, (\beta_0 Z)^{9/2} \gg Z N_p v_{E_0}^2 / v_T^2 \gg 1, \tag{6.1}$$

where the notation

$$N_p = \frac{5\beta_0}{28}\,\frac{\omega_L^2}{c^2}\, Z l l_t(\theta_k) \tag{6.2}$$

has been introduced. The inequalities (6.1) are consistent when not only $Z \gg 1$ but also $N_p \gg 1$ holds. The latter condition can be represented in the form

$$10\beta(\theta_k)\left(\frac{10}{\Lambda}\right)^2\left[\frac{10^{19}}{n[\mathrm{cm}^{-3}]}\right]\left[\frac{T[\mathrm{eV}]}{100}\right]^4 \frac{l_t}{Zl} \gg 1. \tag{6.3}$$

According to Eqs. (6.1) and (6.3) the effect of the nonlocal nature of the transport on radiation filamentation in a turbulent plasma is evident for the typical parameters of a nonisothermal laser plasma. In accordance with (4.11) the spatial scale of the most efficiently amplified filament depends on the turbulent noise level, the electron density and temperature, and also on the radiation intensity I and wavelength $\lambda_0 = 2\pi/k_0 \simeq 2\pi c/\omega_0$ and has the form

$$\begin{aligned}\lambda_m[\mathrm{cm}] &= 2\pi/k_m \\ &= 7.4\cdot 10^{-4}\left(\frac{10}{\Lambda}\right)^{2/9}\left[\frac{T[\mathrm{eV}]}{100}\right]^{5/6}\left[\frac{10^{19}}{n[\mathrm{cm}^{-3}]}\right]^{11/18} \\ &\quad\times\left[\beta(\theta_k)\,\frac{l_t}{Zl}\right]^{1/9}\left[\lambda_0^2[\mu\mathrm{m}]\, I\left[\frac{\mathrm{W}}{\mathrm{cm}^2}\right]10^{-15}\right]^{-7/18},\end{aligned} \tag{6.4}$$

where $I = c|E_0|^2/4\pi$. The scale λ_m is greater than the radiation wavelength λ_0 but less than the characteristic filament growth length

$$L_m = 2\pi/G_m = \frac{2\sqrt{5}}{3}\,\lambda_m^2/\lambda_0 \gg \lambda_m\,. \tag{6.5}$$

Finally, the threshold value of the radiation intensity for filamentation (4.13) depends strongly on the radiation wavelength and is equal to

$$\begin{aligned}I_{\mathrm{th}}\left[\frac{\mathrm{W}}{\mathrm{cm}^2}\right] &\simeq 5.5\cdot 10^{18}\left(\frac{10}{\Lambda}\right)^{4/7}\frac{1}{Z}\left[\beta(\theta_k)\,\frac{l_t}{Zl}\right]^{2/7}\left(\frac{\lambda_0}{L}\right)^{9/7} \\ &\quad\times\left[\frac{10^{19}}{n[\mathrm{cm}^{-3}]}\right]^{11/7}\left[\frac{T[\mathrm{eV}]}{100}\right]^{15/7}[\lambda_0[\mu\mathrm{m}]]^{-32/7},\end{aligned} \tag{6.6}$$

where λ_0, L, l and l_t are measured in the same units. For example, for a plasma with the parameters $n = 10^{19}\ \mathrm{cm}^{-3}$, $T = 300$ eV, $Z = 5$ and for $\beta(\theta_k) l_t/l = 0.1$, $\lambda_0/L = 0.001$ for radiation with $\lambda_0 = 1\ \mu$m the threshold intensity value amounts to $I_{\mathrm{th}} \simeq 6\times 10^{14}\ \mathrm{W/cm^2}$.

Let us mention some experiments that have a bearing on the theory presented above. First of all let us mention a comparatively recent paper[17] in which filamentation was investigated experimentally under conditions such that nonlocal transport is important. The carbon dioxide plasma in Ref. 17 has the parameters $n = 10^{20}\ \mathrm{cm}^{-3}$, $T = 1$ keV, $Z_{\mathrm{eff}} = 3.5$, for which the electron mean free path is equal to $l = 0.005$ cm. The effective wavenumber, determined by the filament size, amounts to[17] $k_m \sim (3-5)\times 10^3\ \mathrm{cm}^{-1}$, which gives $k_m l \sim (15-25)$ and makes it possible to speak of collisionless thermal electrons. Theoretical representations of filamentation, which take account of both the ponderomotive force[1] and nonlocal heat transport,[6] are used to interpret the results in Ref. 17. The need to take nonlocal transport into account is pointed out. The analysis of the experiment of Ref. 17 must obviously also take account of the effect of turbulence on electron transport. The fact of the matter is that in the experiment of Ref. 17 the inhomogeneity scale of the density amounts to $L_n \sim 100\ \mu$m. If it is assumed that the inhomogeneity scale L_T of the electron temperature does not exceed $3L_n$, then, following the theory of ion-acoustic turbulence,[14] one can estimate the degree to which the ion-acoustic instability exceeds threshold as $p \simeq 0.5(v_T/v_s) \times (l/L_T)$. Using the estimate of Ref. 17 for $v_s \simeq 3 \times 10^7$ cm/sec, we find $p \sim 4$. This makes it possible to speak of the possibility of exciting ion-acoustic turbulence by an electron heat flux. Following the theory presented above, we estimate the parameter $N_t \sim 0.1 Z k_m^2 l l_t \sim 0.1(Z/p)(k_m l)^2 \sim 20-50$. Since $N_t \gg 1$ holds, nonlocal heat transport can also appear in the experiment of Ref. 17 in the turbulent state. An indirect proof of the excitation of turbulence could be obtained by conducting an experiment on the apparatus of Ref. 17 in order to detect anisotropy of the filamentation threshold.

A plasma with density $n \simeq 6\times 10^{16}\ \mathrm{cm}^{-3}$ and a low temperature $T \simeq 10$ eV, also heated by CO_2 laser radiation and having a temperature inhomogeneity scale $L_T \simeq 0.01$ cm is also an appropriate object for investigating filamentation in a turbulent plasma.[18,19] In such a plasma, according to Refs. 18, 19, ion-acoustic instability, the excitation threshold of which is exceeded by a factor of more than 10, is the cause of heat-transport restriction. Effects that are most clearly caused by ion-acoustic turbulence are evident in the current-carrying plasma of straight discharges, where an ion-acoustic instability is easily excited by the current.[20] For the estimates let us choose plasma parameters which are already available for these apparatuses: $n \simeq 10^{16}\ \mathrm{cm}^{-3}$, $T \simeq 100$ eV, $Z = 5$, $l_t(\theta_k)/l \simeq 0.1$, $l \simeq 0.3$ cm, $L \simeq 50$ cm. Let us consider the interaction of the radiation of a CO_2 laser ($\lambda_0 = 10\ \mu$m), usually used for investigating turbulent spectra (see, for example, Ref. 21), with such a plasma. Then we obtain an estimate of $I_{\mathrm{th}} = 5\times 10^{11}\ \mathrm{W/cm^2}$ for the filamentation threshold from Eq. (6.6). For such a radiation intensity the size of the most efficiently growing filament (6.4) amounts to $\lambda_m \simeq 0.1$ cm, and its gain length (6.5) $L_m \simeq 15$ cm. In this case the parameter $N_t \sim 50$ and the nonlocal heat transport must be taken into account.

Experiments to detect these filamentation mechanisms are also possible at microwave frequencies. As an example of suitable experimental conditions we can point to those which are close to those implemented in Refs. 22, 23, aimed at an investigation of high-frequency heating and electron transport anomalies. Let us point out that the possibility for the appearance of effects caused by ion-acoustic turbulence has already been mentioned in these experiments.[24]

Let us now discuss the possibility of observing SBS in a weakly collisional plasma under nonlocal electron transport conditions. For greater clarity of the estimates we assume that sound damping by ions is less than by electrons. Moreover, let us also consider the situation when inequalities (5.15) and (3.17) are satisfied simultaneously. In this situation SBS mechanisms are realized that correspond to the simultaneous appearance of a change in sound damping and a renormalization of the effective ponderomotive force because of nonlocal electron transport. In this situation the wavenumber of the acoustic vibrations determining the scattering lies within the interval

$$Z^{5/4}\sqrt{\frac{l}{l_t(\theta_k)}}\,\min\{1,[l/3\sqrt{Z}l_t(\theta_k)]^{7/6}\}\gg kl\gg\sqrt{\frac{l}{Zl_t(\theta_k)}}. \tag{6.7}$$

For such wave numbers we have from Eq. (5.14) the following expression for the radiation intensity corresponding to SBS threshold:

$$I_{\mathrm{th}}\left[\frac{\mathrm{W}}{\mathrm{cm}^2}\right]\simeq 2\cdot 10^{11}\left(\frac{\Lambda}{10}\right)\left(\frac{Z}{A}\right)^{1/2}(Zkl)^{1/7}\lambda_0^{-1}[\mu\mathrm{m}]\times\left[\frac{n[\mathrm{cm}^{-3}]}{10^{19}}\right]\left[\frac{100}{T[\mathrm{eV}]}\right]^{1/2}\left[\frac{Zl}{l_t(\theta_k)}\right]^{3/7}, \tag{6.8}$$

where A is the ratio of the ion and proton masses. In turn, we find

$$G_{-1}[\mathrm{cm}^{-1}]=4.4\cdot 10^{-2}\left(\frac{\Lambda}{10}\right)\times\left[\frac{n[\mathrm{cm}^{-3}]}{10^{19}}\right]^2\left[\frac{100}{T[\mathrm{eV}]}\right]^{3/2}\lambda_0^2[\mu\mathrm{m}]\left(\frac{I}{I_{\mathrm{th}}}-1\right). \tag{6.9}$$

from Eq. (5.19) for the spatial growth rate. If $Z=5$, $A/Z=2$, $n=10^{19}$ cm^{-3}, $T=300$ eV, $\lambda_0=1$ μm, $l/l_t(\theta_k)=10$, then for $kl=5$ we have $I_{\mathrm{th}}=7\times 10^{11}$ W/cm^2 from Eq. (6.8). An estimate of $G_{-1}\simeq 8\times 10^{-3}(I/I_{\mathrm{th}}-1)$ cm^{-1} follows from Eq. (6.9) for these same conditions. The SBS threshold is significantly lower than the filamentation threshold. This means that the characteristics of SBS under nonlocal transport conditions in a turbulent plasma should be evident in the experiments considered previously in discussing the possibility of observing filamentation. Observation of these characteristics would make it possible to obtain experimental information on the turbulent noise level as well as to establish the degree of anisotropy of the ion-acoustic fluctuations of the charge density.

This work was performed under Project No. 97-02-16779 of the Russian Fund for Fundamental Research.

We wish to thank the reviewer for encouraging us to include a more detailed discussion of the experiments.

[1] W. L. Kruer, Plasma Phys. Controlled Fusion **9**, 63 (1985).
[2] B. G. Eremin and A. G. Litvak, JETP Lett. **13**, 430 (1971).
[3] R. D. Jones, W. C. Mead, S. V. Coggeshell, C. H. Aldrich, J. L. Norton, G. D. Pollak, and J. M. Wallace, Phys. Fluids **31**, 1249 (1988).
[4] F. Cornolti and M. Lucchesi, Plasma Phys. Controlled Fusion **31**, 213 (1989).
[5] S. I. Braginskiĭ, *Reviews of Plasma Physics* (Atomizdat, Moscow, 1963), vol. 1 (Consultants Bureau, New York, 1963).
[6] E. M. Epperlein, Phys. Rev. Lett. **65**, 2145 (1990).
[7] A. V. Maksimov and V. P. Silin, Zh. Éksp. Teor. Fiz. **103**, 73 (1993) [JETP **76**, 39 (1993)].
[8] P. K. Shukla, Phys. Fluids B **5**, 4253 (1993).
[9] A. V. Maximov and V. P. Silin, Phys. Lett. A **192**, 67 (1994).
[10] K. N. Ovchinnikov and V. P. Silin, Fiz. Plazmy **22**, 436 (1996) [Plasma Phys. Rep. **22**, 395 (1996)].
[11] K. N. Ovchinnikov and V. P. Silin, Kratk. Soobshch. Fiz. , (1997).
[12] V. P. Silin and S. S. Uryupin, JETP Lett. **63**, 611 (1996).
[13] V. P. Silin and S. A. Uryupin, Zh. Éksp. Teor. Fiz. **110**, 2028 (1996) [JETP **83**, 1118 (1996)].
[14] V. P. Silin, *Spring College on Plasma Physics* (SMR/150-23, Trieste, Italy, 1985), p. 1.
[15] V. P. Silin, Kratk. Soobshch. Fiz. **5**, 59 (1983).
[16] V. P. Silin and S. A. Uryupin, Zh. Éksp. Teor. Fiz. **98**, 117 (1990) [Sov. Phys. JETP **71**, 64 (1990)].
[17] P. E. Young, Phys. Plasmas **2**, 2815 (1995).
[18] D. R. Gray, J. D. Kilkenny, M. S. White, P. Blyth, and D. Hull, Phys. Rev. Lett. **39**, 1270 (1977).
[19] D. R. Gray and J. D. Kilkenny, Plasma Phys. **22**, 81 (1980).
[20] E. D. Volkov, N. F. Perepelkin, V. A. Suprunenko, and E. A. Sukhomlin, *Collective Phenomena in Current-Carrying Plasma* (Naukova Dumka, Kiev, 1978), p. 188.
[21] B. F. M. Pots, J. J. H. Coumans, and D. C. Schram, Phys. Fluids **24**, 517 (1981).
[22] J. M. Liu, J. S. De Groot, J. P. Matte, T. W. Johnson, and R. P. Drake, Phys. Rev. Lett. **72**, 2717 (1994).
[23] J. M. Liu, J. S. De Groot, J. P. Matte, T. W. Johnson, and R. P. Drake, Phys. Plasmas **1**, 3570 (1994).
[24] J. S. De Groot, J. M. Jiu, J. P. Matte, D. Perdue, T. W. Johnston, R. P. Drake, K. G. Estabrook, W. L. Kruer, A. B. Langdon, V. Yu. Bychenkov, V. P. Silin, and S. A. Uryupin, AIP Conf. Proc. (USA) No. **318,** 135 (1994).

Translated by Eugene R. Heath

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Negative viscosity for Rossby wave and drift wave turbulence

A. V. Chechkin, M. I. Kopp, and V. V. Yanovsky

Institute for Single Crystals Acad. Sci. Ukraine, 310001, Kharkov, Ukraine

A. V. Tur

Laboratoire d'Energetique et de Mecanique Theorique et Appliquee, CNRS-URA 875, 2, Avenue de la Foret de Haye-B.P. 160 54504, Vandoeuvre les Nancy Cedex—France

(Submitted 20 February 1997)
Zh. Éksp. Teor. Fiz. **113**, 646–663 (February 1998)

The possible occurrence of a ''negative viscosity effect'' is studied for Rossby wave and drift wave turbulence. It is assumed that (i) the space and time scales of the wave field are much smaller than the scales of the mean field, and (ii) the small-scale field is sufficiently weak, stationary, and maintained by an external source. Such a formulation is fruitful for studying the effects (characterized by the effective viscosity) of smaller-scale motions upon larger-scale ones. The criteria of large-scale instability due to the negative effective viscosity are derived for the coherent wave motions as well as for small-scale isotropic wave turbulence. © *1998 American Institute of Physics.* [S1063-7761(98)02002-2]

1. INTRODUCTION

The processes of pattern formation have been extensively studied in various hydrodynamic models. One of the aspects of this problem has been called ''negative viscosity.'' This term was introduced when analyzing large-scale geophysical experiments; see Refs. 1,2. In the modern literature this term implies two connected classes of phenomena. The first of these is related to the description of anomalous flows of the turbulent kinetic energy through the spectrum toward the region of small wavenumbers in two-dimensional (2D) hydrodynamics and to the formation of stationary turbulent spectra. This problem has been studied in Ref. 3 for 2D homogeneous isotropic turbulence in a Navier–Stokes (NS) fluid with zero mean velocity, and for 2D magnetohydrodynamics in Ref. 4. Using the same closure techniques of the direct-interaction family, the authors show that the negative eddy damping rate occurs for both cases. Another class of phenomena, to which our paper is devoted, is related to pattern formation when the turbulent spectrum is assumed known. Here the negative-viscosity effect means the appearance of a negative dissipative factor in the equation for the mean flow. From the theoretical viewpoint, generation of large-scale structures is understood as a manifestation of long-wavelength instability in a system of small-scale vortices or waves, the energy of small-scale motions being constant (it is mathematically convenient to treat the small-scale motions as generated by an external source).

A number of analytical studies of the negative viscosity effect were initiated by two-dimensional flow of a viscous incompressible fluid, which is undamped due to the existence of an external force periodic along one of the coordinates. In this paper the instability criterion for a sinusoidal velocity profile and the marginal stability curve were derived. Along with this paper, the problem has been considered in Ref. 6 and generalized to an arbitrary periodic velocity profile in Ref. 7.

When studying linear stability and nonlinear regimes appearing, it is convenient to use the two-scale expansion method. In this method it is assumed that the characteristic space and time scales of the basic initial motions are smaller than the scales of the secondary flows. Therefore, it is possible to introduce a small parameter characterizing the ratio of the characteristic scale of small-scale motions to that of large-scale secondary motions. The solution of the hydrodynamic equations is sought in the form of an expansion in the small parameter, while the equation describing the evolution of the large-scale component is obtained from the solvability condition of the initial equations in the corresponding approximation order. Two-scale expansions are widely used in the other problems, which are connected with generation of large-scale fields and structures by small-scale fields and motions. As examples, we mention the papers on the kinetic α-effect[8] and on generation of large-scale convective patterns by helical turbulence.[9]

Using the two-scale formalism, the equations of a weakly nonlinear theory for the large-scale motions have been obtained and studied analytically and numerically for problems with the Kolmogorov flows.[10,11] In the case when the small-scale motions are describable as homogeneous turbulence, negative viscosity effects have been studied in Ref. 12. In particular, it has been shown for 2D NS flows that a homogeneous isotropic small-scale turbulence does not lead to negative eddy viscosity. A general multiscale formalism for the study of eddy viscosities for incompressible flows of arbitrary dimensionality has been developed in Ref. 13. In this paper explicit expressions for eddy viscosity in terms of the correlation function of the small-scale basic flow have been derived for the low Reynolds number isotropic case (in accordance with Ref. 12, eddy viscosity enhances molecular viscosity), and for the parallel time-independent flow, of which Kolmogorov flow is an example. Such parallel flow undergoes a negative viscosity instability with respect to large-scale perturbations transverse to the basic flow. Among the papers close in spirit to this group we also mention our previous work,[14] where the negative viscosity effect was found during the excitation of a single drift wave (with wavelength greater than the ion Larmor radius at electron temperature) in a magnetized inhomogeneous plasma and particular solutions of the weakly nonlinear equation for large-scale perturbations were studied.

In contrast to the papers on liquid hydrodynamics mentioned above, the present paper deals with negative viscosity in Rossby wave turbulence and drift wave turbulence. Rossby turbulence is a widespread type of wave motion in the ocean and atmosphere; see, for example, Refs. 15 and 16. Drift turbulence is widespread in magnetized inhomogeneous plasmas of numerous thermonuclear devices and the ionosphere; see, for example, Refs. 17 and 18. It is well-known that, despite the quite different physical origin of these motions, their formal description is very similar.[19] Moreover, Rossby wave turbulence and drift wave turbulence obey identical nonlinear partial differential equations (in the simplest description). Therefore, it is natural to discuss them both together. In order to clarify the discussion and the results, we use the simplest method of analysis which allows us to elucidate in a uncomplicated way the appearance of nontrivial effects and to find out how they differ from non-wave hydrodynamic problems. We assume that it is possible to divide the fields into a large-scale slowly varying part and a small-scale rapidly evolving part. The small-scale field is a wave field, whose level is kept stationary due to the existence of an external source (external force) in the initial equation. The evolution of the large-scale part is calculated by averaging over the small-scale part. In such a formulation the effective (turbulent) viscosity determining the evolution of the large-scale field is a functional of the given spectrum of waves. We use the simplest model spectra to demonstrate the contrast with the hydrodynamic problems mentioned above. In particular, we show that small-scale isotropic Rossby and drift wave turbulence can act as a negative effective viscosity on large-scale perturbations. This points to a more substantial role of a nonlocal energy transfer from small scales to larger ones in the case of the Rossby and drift wave turbulence than in the case of 2D NS turbulence.

2. EQUATION FOR THE LARGE-SCALE FIELD EVOLUTION

We start from the well-known model two-dimensional equation, which describes the space–time evolution of the stream function in the Rossby wave theory.[20] In dimensionless units,

$$\frac{\partial\psi}{\partial t}-\frac{\partial}{\partial t}\Delta\psi-\frac{\partial\psi}{\partial x}+\nu\Delta^2\psi-[\nabla\psi,\nabla\Delta\psi]_z=F. \tag{2.1}$$

Here $\nabla=\mathbf{e}_x\partial/\partial x+\mathbf{e}_y/\partial/\partial y$, $\Delta=\partial^2/\partial x^2+\partial^2/\partial y^2$, and ν is the (dimensionless) molecular viscosity of the gas or liquid. Following the familiar method, used in the turbulence theory, we introduce a source into the right hand side of Eq. (2.1). The role of this source is to maintain the stationary level of the wave turbulence.

For the Rossby wave the x-axis denotes the latitude direction (from west to east) and the y-axis denotes the meridional one (from south to north). We note that usually for drift waves the x-axis denotes the radial direction in a thermonuclear device, or the direction in which plasma density varies, while the y-axis denotes the azimuthal direction, and so one has to replace $\partial\psi/\partial x$ by $-\partial\psi/\partial y$ in Eq. (2.1) in order to follow the conventional notation used in drift wave theory. Then ψ is the dimensionless potential, and ν is the (dimensionless) ion viscosity for magnetized plasmas. However, in this paper, for definiteness, we use the ‘‘Rossby wave coordinate frame,’’ i.e., that in which Eq. (2.1) is written. Obviously, the final results can be easily reproduced in a ‘‘drift wave coordinate frame.’’

In the linear approximation Eq. (2.1) describes the wave propagation with the frequency

$$\omega_R=-\frac{k_x}{1+k^2}, \tag{2.2}$$

and the damping rate

$$\nu_R=\frac{\nu k^4}{1+k^2}, \tag{2.3}$$

where $\mathbf{k}$ is the wave vector, $k^2=k_x^2+k_y^2$.

Now we divide the field ψ into mean and fluctuating (turbulent) components:

$$\psi=\overline{\psi}+\psi^T; \tag{2.4}$$

the bar denotes statistical averaging and ‘‘T’’ means ‘‘turbulent.’’ After averaging Eq. (2.1) we get

$$\frac{\partial\overline{\psi}}{\partial t}-\frac{\partial}{\partial t}\Delta\overline{\psi}-\frac{\partial\overline{\psi}}{\partial x}+\nu\Delta^2\overline{\psi}$$
$$=[\nabla\overline{\psi}\times\nabla\Delta\overline{\psi}]_z+[\overline{\nabla\psi^T\times\nabla\Delta\psi^T}]_z. \tag{2.5}$$

To get a closed equation for $\overline{\psi}$ it is necessary to express

$$Q=[\overline{\nabla\psi^T\times\nabla\Delta\psi^T}]_z \tag{2.6}$$

in terms of $\overline{\psi}$. The equation for ψ^T is obtained from Eqs. (2.1) and (2.5):

$$\frac{\partial\psi^T}{\partial t}-\frac{\partial}{\partial t}\Delta\psi^T-\frac{\partial\psi^T}{\partial x}+\nu\Delta^2\psi-[\nabla\overline{\psi}\times\nabla\Delta\psi^T]_z-[\nabla\psi^T$$
$$\times\nabla\Delta\overline{\psi}]_z-([\nabla\psi^T\times\nabla\Delta\psi^T]_z-[\overline{\nabla\psi^T\times\nabla\Delta\psi^T}]_z). \tag{2.7}$$

Hence the closed equation for $\overline{\psi}$ implies the application of some closure procedure. Since we are interested here in negative viscosity effects for large-scale flows, we use the following approach. Assume that the mean quantities vary on space and time scales which are larger than the characteristic scales of the fluctuation fields. We introduce the characteristic size l of the small-scale field and the characteristic size L of the large-scale field. Then we estimate the ratio of the quantities $\nu\Delta^2\psi^T,[\nabla\overline{\psi}\times\nabla\Delta\psi^T]_z$, $[\nabla\psi^T\times\nabla\Delta\overline{\psi}]_z$, $[\nabla\psi^T\times\nabla\Delta\psi^T]_z$. They stand in the ratios

$$1:\frac{l}{L}\,\mathrm{Re}_L:\frac{l^3}{L^3}\,\mathrm{Re}_L:\mathrm{Re}_l\,,$$

where $\mathrm{Re}_L\approx\overline{V}L/\nu$ is the Reynolds number of the large-scale motions and $\mathrm{Re}\approx V^T l/\nu$ is the Reynolds number of the small-scale ones (V and V^T are characteristic velocities of large-scale and small-scale motions, respectively). Therefore, for sufficiently small Re_l we can neglect terms in Eq. (2.7) which are quadratic in ψ^T. Furthermore, in accordance with multiscale expansion schemes, we introduce the ''slow'' variable $\mathbf{X}$ and the fast variable $\mathbf{x}$. The average quantities depend on the slow variable only, while the fluctuating components depend on both the fast and slow variables. The following inequality holds:

$$\left|\frac{\partial}{\partial\mathbf{X}}\right|\approx|\mathbf{K}|\ll\left|\frac{\partial}{\partial\mathbf{x}}\right|\approx|\mathbf{k}|, \tag{2.8}$$

where $\mathbf{K}$ and $\mathbf{k}$ are large-scale and small-scale wave vectors, respectively.

Thus, we can find the solution for ψ^T as an expansion in powers of K, that is,

$$\psi^T=\psi^{(0)}(\mathbf{x},t)+\psi^{(1)}(\mathbf{x},\mathbf{X},t)+\ldots+\psi^{(4)}(\mathbf{x},\mathbf{X},t). \tag{2.9}$$

The solution to order K^4 is presented in Appendix A. Then, the functional dependence of Q through $\overline{\psi}$ is obtained there. So instead of Eq. (2.5) we get the following equation for the large-scale part of ψ:

$$\begin{aligned}\hat{L}\overline{\psi}=&\left(\frac{\partial\overline{\psi}}{\partial X}\frac{\partial}{\partial Y}\Delta_s\overline{\psi}-\frac{\partial\overline{\psi}}{\partial Y}\frac{\partial}{\partial X}\Delta_s\overline{\psi}\right)\\&+\varepsilon_{mn}\varepsilon_{jk}\Bigg[\eta_{mpk}\frac{\partial^3\overline{\psi}}{\partial X_p\partial X_n\partial X_j}-\nu^{(1)}_{mk}\frac{\partial^2\Delta_s\overline{\psi}}{\partial X_n\partial X_j}\\&+\nu^{(2)}_{mpk}\frac{\partial}{\partial X}\left(\frac{\partial^3\overline{\psi}}{\partial X_p\partial X_n\partial X_j}\right)+\nu^{(3)}_{mlpk}\frac{\partial^4\overline{\psi}}{\partial X_l\partial X_n\partial X_p\partial X_j}\\&+\varepsilon_{rq}N^{(1)}_{mlkq}\frac{\partial^2}{\partial X_l\partial X_n}\left(\frac{\partial\overline{\psi}}{\partial X_r}\frac{\partial\overline{\psi}}{\partial X_j}\right)\\&+\varepsilon_{rq}N^{(2)}_{kpmq}\frac{\partial\overline{\psi}}{\partial X_j}\frac{\partial^3\overline{\psi}}{\partial X_p\partial X_n\partial X_r}\\&-\varepsilon_{rq}N^{(3)}_{kpnq}\frac{\partial^2\overline{\psi}}{\partial X_m\partial X_j}\frac{\partial^2\overline{\psi}}{\partial X_p\partial X_r}\Bigg],\end{aligned} \tag{2.10}$$

where

$$\hat{L}=\frac{\partial}{\partial t}-\frac{\partial}{\partial t}\Delta_s-\frac{\partial}{\partial X}+\nu\Delta_s^2. \tag{2.11}$$

$$\Delta_s\equiv\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2}. \tag{2.12}$$

$(j,l,m,n,p,q,r)=x,y,\varepsilon_{mn}$ is the unit second-rank antisymmetric tensor,

$$\varepsilon_{11}=\varepsilon_{22}=0,\quad \varepsilon_{12}=-\varepsilon_{21}=1,\quad (m,n)=1,2,$$

$$\begin{aligned}\eta_{mpk}&=\int\int d\mathbf{k}d\omega\Pi_1(k,\omega)2k_mk_pk_kk^2(\omega-\omega_R),\\ \nu^{(1)}_{mk}&=\int\int d\mathbf{k}d\omega\Pi_1(k,\omega)\nu_Rk_mk_kk^2,\\ \nu^{(2)}_{mpk}&=\int\int d\mathbf{k}d\omega\Pi_2(k,\omega)4k_mk_pk_kk^2\nu_R(\omega-\omega_R).\\ \nu^{(3)}_{mlpk}&=\int\int d\mathbf{k}d\omega\Pi_2(k,\omega)8k_mk_lk_pk_kk^2\\&\quad\times[\nu k^2(\nu_R^2-(\omega-\omega_R)^2)+\omega\nu_R(\omega-\omega_R)],\\ N^{(1)}_{mlkq}&=\int\int d\mathbf{k}d\omega\Pi_2(k,\omega)2k_mk_lk_qk_kk^2\\&\quad\times[(\omega-\omega_R)^2-\nu_R^2],\\ N^{(2)}_{kpmq}&=\int\int d\mathbf{k}d\omega\Pi_2(k,\omega)2k_kk_pk_mk_qk^4(k^2+1)^{-1},\end{aligned} \tag{2.13}$$

$$\begin{aligned}\Pi_1(\mathbf{k},\omega)&=\frac{\Phi(\mathbf{k},\omega)}{(1+k^2)[(\omega-\omega_R)^2+\nu_R^2]},\\ \Pi_2(\mathbf{k},\omega)&=\frac{\Phi(\mathbf{k},\omega)}{(1+k^2)^2[(\omega-\omega_R)^2+\nu_R^2]^2},\end{aligned} \tag{2.14}$$

and $\Phi(\mathbf{k},\omega)$ is the space–time spectral density for the small-scale field.

Let us discuss the meaning of the terms in Eqs. (2.10), (2.13). The first term in Q gives the correction to the frequency in the dispersion equation for the large-scale motions. The next three terms in Q are the ones which lead either to damping (positive effective viscosity) or to growth (negative effective viscosity) of the large-scale motions. The last three terms describe the nonlinear interaction of the large-scale motions. In this paper we are interested mainly in terms of the viscous type. We consider the effects depending on the properties of small-scale wave turbulence. However, in order to make the discussion simpler and to clarify the differences between our paper and the papers mentioned above, in the next section we demonstrate for the case of the 2D flow of a viscous incompressible fluid what effect can be responsible for the appearance of the negative effective (turbulent) viscosity.

3. QUALITATIVE CONSIDERATION OF THE ORIGIN OF THE NEGATIVE VISCOSITY TERM

In order to simplify the discussion as much as possible, we consider here the equation for the stream function of a 2D incompressible viscous fluid:[21]

$$\frac{\partial}{\partial t}\Delta\psi-\nu\Delta^2\psi+[\nabla\psi\times\nabla\Delta\psi]_z=F. \tag{3.1}$$

Inserting ψ in the form (2.4), we get equations analogous to Eqs. (2.5)–(2.7):

$$\frac{\partial}{\partial t}\Delta\overline{\psi}-\nu\Delta^2\overline{\psi}-\overline{[\nabla\psi^T\times\nabla\Delta\psi^T]_z}=0, \tag{3.2}$$

$$\frac{\partial}{\partial t}\Delta\psi^T-\nu\Delta^2\psi^T+[\nabla\overline{\psi}\times\nabla\Delta\psi^T]_z+[\nabla\psi^T\times\nabla\Delta\overline{\psi}]_z=F. \tag{3.3}$$

It is noteworthy that the third and fourth terms on the left-hand side of Eq. (3.3) describe the interaction between small- and large-scale fields. The term $[\nabla \overline{\psi}\times\nabla\Delta\psi^T]_z$ describes the transport of the fluctuation vorticity $W_z^T=\Delta\psi^T$ by the mean flow $\langle\mathbf{V}\rangle=[\mathbf{e}_z\times\nabla\overline{\psi}]$, while the term $[\nabla\psi^T\times\nabla\Delta\overline{\psi}]_z$ describes the transport of the mean vorticity by the fluctuation component. As in Sec. 2, we introduce the natural physical assumption that the space and time scales of the average quantities are larger than the scales of the fluctuations. Introducing fast and slow variables (x and X, respectively), we use the Fourier transform over the fast variable:

$$\psi^T(\mathbf{x},t)=\int\frac{d\mathbf{k}d\omega}{(2\pi)^3}\,\hat{\psi}^T(\mathbf{k},\omega)\exp(-i\omega t+i\mathbf{k}\cdot\mathbf{x}).$$

We note that in the Fourier representation the term $[\nabla\overline{\psi}\times\nabla\Delta\psi^T]_z$ describes the Doppler shift of the fluctuation frequency. From Eq. (3.3) we get

$$\hat{\psi}^T(\mathbf{k},\omega)\approx\frac{\hat{F}(\mathbf{k},\omega)}{k^2(i\omega-\nu k^2)}\left\{1+\frac{1}{i\omega-\nu k^2}\left(ik_n\varepsilon_{mn}\frac{\partial\overline{\psi}}{\partial X_m}-\frac{ik_m}{k^2}\varepsilon_{mn}\frac{\partial^3\overline{\psi}}{\partial X_n^3}\right)\right\}. \tag{3.4}$$

where it is assumed for simplicity that the terms containing slow spatial derivatives are small. We note once more that the interaction of the fluctuations ψ^T with the mean flow $\overline{\psi}$ (in Eq. (3.4) the terms with the first and third slow derivatives are due to this interaction) causes ψ^T to depend on the slow coordinate. This fact implies that among the terms entering into $-\langle[\nabla\psi^T\times\nabla\Delta\psi^T]_z\rangle$ the following term occures:

$$\overline{\varepsilon_{mn}\frac{\partial\psi^T}{\partial x_m}\frac{\partial}{\partial X_n}\Delta_s\psi^T}, \tag{3.5}$$

where Δ_s is the slow Laplacian, as before, and Δ has to be replaced by Δ_s in the first two terms in Eq. (3.2).

Inserting Eq. (3.4) into Eq. (3.5), we can see that Eq. (3.5) gives rise to the following terms in the left-hand side of Eq. (3.2):

$$-\varepsilon_{mn}\varepsilon_{jk}\nu_{mk}^{(1)}\frac{\partial^2}{\partial X_n\partial X_j}\Delta_s\overline{\psi}+\varepsilon_{mn}\varepsilon_{jk}\mu_{mj}\frac{\partial^4}{\partial X_n\partial X_k^3}\Delta_s\overline{\psi}, \tag{3.6}$$

$$\nu_{mk}^{(1)}=\int d\mathbf{k}d\omega\langle\hat{F}^2\rangle_{\mathbf{k},\omega}\frac{\nu k^2}{k^4(\omega^2+\nu^2k^4)^2}k_mk_k, \tag{3.7}$$

$$\mu_{mj}=\int d\mathbf{k}d\omega\langle\hat{F}^2\rangle_{\mathbf{k},\omega}\frac{\nu k^2}{k^6(\omega^2+\nu^2k^4)^2}k_mk_j. \tag{3.8}$$

The first term in Eq. (3.6) is due to the transport of the fluctuation vorticity by the mean flow, while the second term is caused by the transport of the mean vorticity by the fluctuation velocity. The first term corresponds to $\nu_{mk}^{(1)}$, which is calculated for the Rossby wave turbulence in Sec. 2; see Eqs. (2.13). It can be seen that even for isotropic fluctuations the former effect makes a negative contribution to the turbulent viscosity. Indeed, let $(\hat{F}^2)_{\mathbf{k},\omega}$ be an isotropic function of k. Then,

$$\nu_{mk}^{(1)}=\nu^{(1)}\delta_{mk},\quad \mu_{mj}=\mu\delta_{mj},\quad \nu^{(1)},\mu>0$$

and Eq. (2.6) is rewritten in the form

$$\nu^{(1)}\Delta_s^2\overline{\psi}+\mu\left(\frac{\partial^4}{\partial X^4}+\frac{\partial^4}{\partial Y^4}\right)\Delta_s\overline{\psi}. \tag{3.9}$$

Therefore, the term with $\nu^{(1)}$ gives a negative contribution to the turbulent viscosity. The term with μ describes the dissipation of the large-scale component and bounds the instability region for small wavenumbers.

Thus, the interaction between the large-scale flow and small-scale velocity fluctuations, which manifests itself mainly in the transport of the small-scale fluctuations by the mean flow, gives rise to a viscous term with a negative viscosity coefficient in the equation for the mean component. We stress that this conclusion is valid both for the 2D model of viscous incompressible fluid, see Eq. (3.1), and for the wave model; see Eq. (2.1).

The above treatment is obviously incomplete: it only points to the mechanism for the appearance of the viscous terms with negative viscosity coefficient in the equation for the mean flow. A detailed consideration demands more accurate analysis of Eqs. (3.2) and (3.3) with the two-scale dependence of fluctuations taken into account. The complementary viscous terms can suppress the negative contribution, which is controlled by the transport of the turbulent fluctuations by the large-scale flow. In order to compare our results with those of other authors studying turbulent viscosity by means of the two-scale expansion, and to call attention to the differences between the turbulent viscosity for the 2D NS flows and that for the Rossby and drift waves, in Appendix B we obtain and analyze the equation for the average stream function of 2D incompressible viscous fluid [see Eq. (B1)]. It is a limiting case of Eq. (2.10) if we neglect dispersion ($1+k^2\to k^2$) and the eigenfrequency ($\omega_R=0$) of the waves. In Appendix B we demonstrate that our results correspond to those of Refs. 12 and 13 for 2D NS flows. We also demonstrate that isotropic small-scale fluctuations do not give rise to a negative viscosity in the framework of 2D NS equations. The negative contribution to the eddy viscosity given by $\nu^{(1)}$ in Eq. (3.9) is balanced by the positive contribution. However, in the next Section we demonstrate that this is not the case for the isotropic wave turbulence.

4. NEGATIVE VISCOSITY FOR MODEL SPECTRA OF THE WAVE TURBULENCE

Let us explore the consequences of the general expressions (2.10)–(2.14). We consider the problems arising here, taking the tensor $\nu_{mk}^{(1)}$ as the way of example; see Eqs. (2.13). This tensor is a functional of the space–time spectral function $\Phi(\mathbf{k},\omega)$ of a given small-scale field. The spectrum has a peak on ω at $\omega\approx\omega_R$ and some characteristic width γ_k. The spectrum is multiplied by the Lorentzian curve

$$\frac{\nu_R}{(\omega-\omega_R)^2+\nu_R^2},$$

and integrated over ω and $\mathbf{k}$. Obviously, the result of integration over ω depends on the ratio between the characteristic widths of the multiplied functions, that is, γ_k and ν_R, whereas the result of integration over $\mathbf{k}$ depends on the degree of spectral anisotropy in $\mathbf{k}$. Under these circumstances, it is natural to consider first the simplest model wave spectra leading to the negative viscosity. Thus, we consider the following examples:

1. A small-scale coherent field, that is, the frequency and wavenumber spectra are narrower than the other characteristic widths of the problem.

2. A small-scale isotropic turbulent field. Such a field can be formed in a small-scale region ($k \gg 1$) due to mode coupling, which isotropizes the spectrum (see, e.g., Ref. 22).

4.1. Coherent wave spectrum

We take the wave spectrum as follows:

$$\begin{aligned}\Phi(\mathbf{k},\omega)=\frac{\langle\psi^2\rangle}{2}\,[\delta(\omega-\omega_0)\,\delta(\mathbf{k}-\mathbf{k}_0)\\ +\delta(\omega+\omega_0)\,\delta(\mathbf{k}+\mathbf{k}_0)],\end{aligned}\tag{4.1}$$

where $\langle\psi^2\rangle$ is the variance of the fluctuations,

$$\omega_0(\mathbf{k}_0)=-\frac{k_{0x}}{1+k_0^2}.$$

This case implies that the frequency spectrum is sufficiently narrow:

$$\gamma_{k_0}\ll\nu_R(k_0).\tag{4.2}$$

Inserting Eq. (4.1) into Eqs. (2.10)–(2.14), we get the following equation for $\overline{\psi}$:

$$\begin{aligned}\hat{L}\,\overline{\psi}=&\left(\frac{\partial\overline{\psi}}{\partial X}\,\frac{\partial}{\partial Y}\,\Delta_s\overline{\psi}-\frac{\partial\overline{\psi}}{\partial Y}\,\frac{\partial}{\partial X}\,\Delta_s\overline{\psi}\right)\\
&+\frac{A^2}{2}\,\nu^{-1}k_0^{-2}\left[k_{0y}^2\,\frac{\partial^4\overline{\psi}}{\partial X^4}-2k_{0x}k_{0y}\,\frac{\partial^4\overline{\psi}}{\partial X^3\partial Y}\right.\\
&\left.-2k_{0x}k_{0y}\,\frac{\partial^4\overline{\psi}}{\partial Y^3\partial X}+k_0^2\,\frac{\partial^4\overline{\psi}}{\partial X^2\partial Y^2}+k_{0x}^2\,\frac{\partial^4\overline{\psi}}{\partial Y^4}\right]\\
&+4A^2\nu^{-1}k_0^{-4}\left[-k_{0x}^2k_{0y}^2\,\frac{\partial^4\overline{\psi}}{\partial X^4}+2k_{0x}k_{0y}(k_{0x}^2\right.\\
&-k_{0y}^2)\,\frac{\partial^4\overline{\psi}}{\partial X^3\partial Y}+2k_{0x}k_{0y}(k_{0y}^2-k_{0x}^2)\,\frac{\partial^4\overline{\psi}}{\partial Y^3\partial X}\\
&\left.+(4k_{0x}^2k_{0y}^2-k_{0x}^4-k_{0y}^4)\,\frac{\partial^4\overline{\psi}}{\partial X^2\partial Y^2}-k_{0x}^2k_{0y}^2\,\frac{\partial^4\overline{\psi}}{\partial Y^4}\right]\\
&-A^2\nu^{-2}k_0^{-6}\left[k_{0x}k_{0y}\left(\frac{\partial^2}{\partial Y^2}-\frac{\partial^2}{\partial X^2}\right)+(k_{0x}^2\right.\\
&\left.-k_{0y}^2)\,\frac{\partial^2}{\partial X\partial Y}\right]\left(k_{0x}\,\frac{\partial\overline{\psi}}{\partial Y}-k_{0y}\,\frac{\partial\overline{\psi}}{\partial X}\right)^2\\
&+\frac{A^2}{2}\,\nu^{-2}k_0^{-4}\left[k_{0x}k_{0y}\left(\frac{\partial^2\overline{\psi}}{\partial Y^2}-\frac{\partial^2\overline{\psi}}{\partial X^2}\right)+(k_{0x}^2\right.\\
&\left.-k_{0y}^2)\,\frac{\partial^2\overline{\psi}}{\partial X\partial Y}\right]\left(k_{0y}\,\frac{\partial}{\partial X}-k_{0x}\,\frac{\partial}{\partial Y}\right)^2\overline{\psi}\\
&+A^2\nu^{-2}k_0^{-4}\left(k_{0y}\,\frac{\partial}{\partial X}-k_{0x}\,\frac{\partial}{\partial Y}\right)\left[k_{0x}k_{0y}\left(\frac{\partial^2\overline{\psi}}{\partial Y^2}\right.\right.\\
&\left.\left.-\frac{\partial^2\overline{\psi}}{\partial X^2}\right)+(k_{0x}^2-k_{0y}^2)\,\frac{\partial^2\overline{\psi}}{\partial X\partial Y}\right]\left(k_{0y}\,\frac{\partial\overline{\psi}}{\partial X}-k_{0x}\,\frac{\partial\overline{\psi}}{\partial Y}\right).\end{aligned}\tag{4.3}$$

For simplicity we consider the case when the small-scale field propagates along the x-axis, $\mathbf{k}_0=k_0\mathbf{e}_x$. Taking $\overline{\psi}$ in the form

$$\overline{\psi}=\exp(-i\Omega t+i\mathbf{K}\cdot\mathbf{X}),$$

we get a linear dispersion equation for the large-scale perturbations:

$$\begin{aligned}\Omega=\frac{-K_x}{1+K^2}-\frac{i}{1+K^2}\left[\nu K_x^4+\left(\frac{7\langle\psi^2\rangle}{\nu}+2\nu\right)K_x^2K_y^2\right.\\ \left.-\left(\frac{\langle\psi^2\rangle}{\nu}-\nu\right)K_y^4\right].\end{aligned}\tag{4.4}$$

It is worth noting that this equation describes not only the waves but also the large-scale structures. Indeed, it follows from Eq. (4.4) that the most ''dangerous'' (that is, rapidly growing) are those large-scale perturbations, which are perpendicular to the direction of the wave propagation, that is, $K_x=0$ and $\mathrm{Re}\,\Omega=0$. This implies that the negative viscosity effect leads not only to nonlocal energy transfer from small-scale waves to large-scale waves, but also to the generation of stationary structures highly elongated along one of the coordinates. We return once more to the discussion of these possibilities below. The growth rate of such perturbations is

$$\mathrm{Im}\ \Omega=\left(\frac{\langle\psi^2\rangle}{\nu}-\nu\right)\frac{K_y^4}{1+K_y^2}.\tag{4.5}$$

$\mathrm{Im}\,\Omega>0$ if $\langle\psi^2\rangle>\nu^2$. We stress that in this case we cannot take $\nu\to 0$ because such a limit is in contradiction with the inequality (4.2). For ''the drift wave coordinate frame'' one has to replace K_y by K_x in Eq. (4.5).

The effect described is analogous to the Kolmogorov flow instability of a 2D viscous incompressible fluid.[5] For a single drift wave with wavelength greater than the ion Larmor radius at electron temperature in a magnetized plasma this effect has been studied in Ref. 14. Nonlinear stationary structures formed due to the negative viscosity effect have been also studied there.

4.2. Isotropic wave spectrum

We consider the case when the wave spectrum is isotropic in $\mathbf{k}$. This case is close to that considered in Appendix B, so below we compare the equations of the Appendix with those obtained in this section. Here the model Lorentzian time spectrum is used:

$$\Omega(\mathbf{k},\omega)=\frac{1}{\pi}\,\frac{\gamma_k}{(\omega-\omega_R)^2+\gamma_k^2}\,\Phi(k).\tag{4.6}$$

The width γ_k is a complicated function of $\Phi(k)$, but its explicit form is not discussed here. It should be stressed that we choose the Lorentzian shape for convenience only. It can be easily verified that the result is not changed qualitatively if one chooses other shapes, for example, a Gaussian or in the form of a step; see Eqs. (B6) where $\omega\rightarrow\omega-\omega_R$.

Introducing Eq. (4.6) into Eqs. (2.13), taking the integrals over ω and using the subsidiary integrals (B7), we arrive at Eqs. (B8) for the viscous terms, where instead of Eqs. (B9) we get

$$\nu^{(1)}=\pi\int dk\ k^3\Phi(k)\ \frac{k^2}{1+k^2}\ \frac{1}{\nu_R+\gamma_k},$$

$$\nu^{(3)}=\pi\int dk\ k^3\Phi(k)\ \frac{k^4}{(1+k^2)^2}\ \frac{2\nu k^2+\gamma_k}{(\nu_R+\gamma_k)^2}. \tag{4.7}$$

It is worthwhile to note that if we set $1+k^2\rightarrow k^2$ (no dispersion), then, we naturally get Eqs. (B9) from Eqs. (4.7). As before, $\nu^{(1)}$ and $\nu^{(3)}$ give negative and positive contributions, respectively, to the effective viscosity. However, due to the wave character of the small-scale perturbations the balance between the two contributions is changed. Indeed, instead of Eq. (B10) we get

$$\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial t}\Delta_s-\frac{\partial}{\partial X}\right)\overline{\psi}+\nu_{\mathrm{eff}}\Delta_s^2\overline{\psi}=(1-P)\left(\frac{\partial\overline{\psi}}{\partial X}\ \frac{\partial}{\partial Y}\ \Delta_s\overline{\psi}\right.$$
$$\left.-\frac{\partial\overline{\psi}}{\partial Y}\ \frac{\partial}{\partial X}\ \Delta_s\overline{\psi}\right), \tag{4.8}$$

where

$$\nu_{\mathrm{eff}}=\nu-\nu^{(1)}+\nu^{(3)}$$
$$=\nu+\pi\int dk\Phi(k)\ \frac{k^5}{(1+k^2)^2}\ \frac{\nu k^4-\gamma_k}{(\nu_R+\gamma_k)^2},$$

$$P=\pi\int_0^\infty dk\ \frac{k^5\Phi(k)[(1+k^2)\gamma_k+\nu k^2(k^2-2)]}{(1+k^2)^2(\gamma_k+\nu_R)^2}. \tag{4.9}$$

For the case of 2D NS flows $\nu_{\mathrm{eff}}>0$ holds for any ratio between the spectral width and νk^2; see Eq. (B11). In contrast, as we can see, two possibilities exist for the Rossby waves.

The first possibility is characterized by $\gamma_k\ll\nu_R$ (narrow isotropic spectrum). This case corresponds to the one when the wave intensity is sufficiently low that the broadening of the spectral lines γ_k is small in comparison not only with ω_R (a widely used definition of weak turbulence), but also with ν_R, which, in turn, is less than ω_R (the damping rate is less than the eigenfrequency). Then, setting $\gamma_k\rightarrow 0$, we get

$$\nu_{\mathrm{eff}}=\nu+\frac{\langle\psi^2\rangle}{2\nu}. \tag{4.10}$$

This result coincides with Eq. (B12) and points to the absence of nonlocal energy transfer to the large-scale region.

The second possibility is realized in the limit $\gamma_k\gg\nu_R$ (broad isotropic spectrum). This inequality implies that the intensity of the waves is greater than in the previous case. However, due to the smallness of the damping rate we may still retain the framework of weak turbulence. In this case from Eq. (4.9) one gets

$$\nu_{\mathrm{eff}}=\nu-\pi\int_0^\infty dk\Phi(k)\ \frac{k^5}{(1+k^2)^2\gamma_k}. \tag{4.11}$$

and ν_{eff} can be negative for a sufficiently high Rossby wave level. We may roughly estimate the criterion for the negative viscosity as

$$\langle\psi^2\rangle>\nu\gamma, \tag{4.12}$$

where γ is some effective spectral line broadening. It resembles the analogous criterion for the coherent waves; see Eq. (4.5).

The evolution equation for the large-scale perturbations has the form (4.8), where

$$P=\pi\int_0^\infty dk\Phi(k)\ \frac{k^5}{(1+k^2)\gamma_k}.$$

Looking for $\overline{\psi}$ in the form

$$\overline{\psi}=\exp(-i\Omega t+i\mathbf{K}\cdot\mathbf{X}), \tag{4.13}$$

and inserting Eq. (4.13) into Eq. (4.8), we get the following dispersion equation for the large-scale perturbations:

$$\Omega=-\frac{K_x}{1+K^2}-i\ \frac{\nu_{\mathrm{eff}}K^4}{1+K^2}. \tag{4.14}$$

Equation (4.14) describes two types of motion. The first one ($K_x\neq 0$) is the large-scale wave, while the second one ($K_x=0$) is the large-scale stationary structure, Re $\Omega=0$. Therefore, if $\nu_{\mathrm{eff}}<0$ holds, two possibilities exist for the energy flow from the small-scale region. The first one is nonlocal energy transfer from the small-scale waves to the large-scale ones. The large-scale waves grow, and it can happen that the spectral gap between the two wave regions disappears, the two-scale approximation is violated, and the turbulence becomes nonstationary. We note, however, that this circumstance does not invalidate our treatment, because we consider the initial stage of instability only. Another possibility is related to energy flow from small-scale waves to large-scale stationary structures highly elongated along one of the coordinates. We remind that as in Sec. 4.1 for ''the drift wave coordinate frame'' one has to replace K_x by $-K_y$ in Eq. (4.14).

In this theory there are no limits on the growth rate of the large scale instability as K increases. Such restrictions can be obtained by taking into account the higher-order terms in the expansion in powers of K; see Sec. 3. This procedure has been carried out for a single drift wave in a magnetized plasma in Ref. 14. It has been shown that the terms of order K^6 in the growth rate of the large-scale instability lead to damping of perturbations in the range $K>K_{\mathrm{max}}$ and to the appearance of a maximum of the growth rate for small K.

Since the large-scale perturbations grow due to the negative viscosity effect, the nonlinear term in Eq. (4.8) becomes important. For the subsequent investigation of large-scale structures it is necessary to analyze the nonlinear equation.

5. RESULTS

Here we have studied new effects of generation of large-scale structures. These appear for Rossby wave turbulence in atmosphere and ocean and for drift wave turbulence in magnetized plasmas. The physical reason for their appearance is related to the change in the sign of the effective (turbulent) viscosity in large-scale motions of the medium (negative viscosity). Therefore, the damping of large-scale motions is replaced by growth, which has to be limited due to nonlinear effects. The small-scale field is stationary and maintained by an external source. Such a formulation is fruitful for studying the effects (characterized by the effective viscosity) of smaller-scale motions upon the larger-scale motion. The results obtained are as follows:

1. With the scope of the two-scale expansion, an equation is obtained which describes the evolution of the mean stream function in the presence of the small-scale Rossby-wave field or the evolution of the mean potential in the presence of the small-scale drift-wave field. General expressions are obtained for the terms describing the influence of small-scale motion, namely, the viscous terms, the dispersion term and the terms nonlinear in the large-scale field. These expressions allow one to study the evolution of large-scale motions, with the assumption that the spectrum of the stationary small-scale field is known.
2. The results obtained admit a transition to the hydrodynamics of a viscous incompressible fluid. The previously known results on the eddy viscosity of the small-scale fluid motions are recovered.
3. The qualitative reason for the negative viscosity effect to appear is the transport of small-scale vorticity by the mean flow. This effect leads to the negative-viscosity contribution to the effective viscosity governing the large-scale motions.
4. It is shown that the coherent wave motions lead to a negative effective viscosity. The criterion of large-scale instability due to the negative viscosity effect is derived.
5. In contrast to the case of the small-scale isotropic motions of a viscous incompressible fluid, small-scale isotropic Rossby wave and drift wave motions can lead to a negative effective viscosity. It is demonstrated that the effective viscosity can be negative if the spectral line broadening is greater than the linear damping rate.

This work has been done in the framework of the ''Structure'' Project, which is financed by the National Academy of Sciences of the Ukraine.

APPENDIX A

Derivation of *Q*

In this Appendix the term Q in Eq. (2.6) is derived. We introduce ''slow'' variables together with the ''fast'' ones. The spatial operators are now written in the form:

$$\frac{\partial}{\partial \mathbf{x}} \to \frac{\partial}{\partial \mathbf{x}} + \frac{\partial}{\partial \mathbf{X}},$$

$$\Delta \to \Delta_\perp + 2\frac{\partial^2}{\partial x_p \partial X_p} + \Delta_s,$$

$$\Delta^2 \to \Delta_\perp^2 + 2\Delta_\perp\Delta_s + 4\frac{\partial^2}{\partial x_p \partial X_p}\Delta_\perp + 4\frac{\partial^2}{\partial x_p \partial X_p}\Delta_s + \Delta_s^2 + 4\left(\frac{\partial^2}{\partial x_p \partial X_p}\right)^2. \tag{A1}$$

Here

$$\Delta_s \equiv \frac{\partial^2}{\partial X_p \partial X_p}, \quad \Delta_\perp = \frac{\partial^2}{\partial x_p \partial x_p}.$$

Then, according to Eqs. (2.9), the term Q can be written in the form

$$\begin{aligned} Q(x,X,t) = {} & q^{(01)} + q^{(02)} + q^{(03)} + q^{(04)} + q^{(10)} + q^{(20)} \\ & + q^{(30)} + q^{(40)} + q^{(11)} + q^{(12)} + q^{(13)} + q^{(21)} \\ & + q^{(22)} + q^{(31)} + O(K^5, K^6, \dots), \end{aligned} \tag{A2}$$

where

$$q^{(00)} = \varepsilon_{mn}\left\langle \frac{\partial \psi^{(0)}}{\partial x_m}\frac{\partial}{\partial x_n}\Delta_\perp \psi^{(0)}\right\rangle = 0$$

due to the homogeneity of the turbulence,

$$\begin{aligned} q^{(01)} = {} & \varepsilon_{mn}\left\langle \frac{\partial \psi^{(0)}}{\partial x_m}\left(\frac{\partial^3}{\partial x_n \partial x_p^2} + 2\frac{\partial^3}{\partial x_n \partial x_p \partial X_p} + \frac{\partial^3}{\partial x_n \partial X_p^2}\right.\right. \\ & \left.\left. + \frac{\partial^3}{\partial x_p^2 \partial X_n} + 2\frac{\partial^3}{\partial x_p \partial X_p \partial X_n} + \frac{\partial^3}{\partial X_p^2 \partial X_n}\right)\psi^{(1)}\right\rangle, \end{aligned}$$

$$q^{(10)} = \varepsilon_{mn}\left\langle \left(\frac{\partial}{\partial x_m} + \frac{\partial}{\partial X_m}\right)\psi^{(1)}\frac{\partial}{\partial x_n}\Delta_\perp \psi^{(0)}\right\rangle. \tag{A3}$$

The remaining terms in Eq. (A2) have similar structure. We retain in Eq. (A2) only those terms which are of order $K^1, \dots, K^4$. As will be seen below it is just these terms which give rise to the negative viscosity effect.

To calculate the terms in Eq. (A2) it is necessary to get expressions for $\psi^{(0)}, \psi^{(1)}, \dots, \psi^{(4)}$. Using Eqs. (A1) and (27) one finds the the following equations:

$$O(K^0){:}\ \frac{\partial \psi^{(0)}}{\partial t} - \frac{\partial}{\partial t}\Delta_\perp \psi^{(0)} - \frac{\partial \psi^{(0)}}{\partial x} + \nu\Delta_\perp^2 \psi^{(0)} = F, \tag{A4}$$

$$\begin{aligned} O(K^1){:}\ & \frac{\partial \psi^{(1)}}{\partial t} - \frac{\partial}{\partial t}\Delta_\perp \psi^{(1)} - \frac{\partial \psi^{(1)}}{\partial x} + \nu\Delta_\perp^2 \psi^{(1)} \\ & - \varepsilon_{jk}\frac{\partial \overline{\psi}}{\partial X_j}\frac{\partial}{\partial x_K}\Delta_\perp \psi^{(0)} = 0, \end{aligned} \tag{A5}$$

$$\begin{aligned} O(K^2){:}\ & \frac{\partial \psi^{(2)}}{\partial t} - \frac{\partial}{\partial t}\Delta_\perp \psi^{(2)} - \frac{\partial \psi^{(2)}}{\partial x} + \nu\Delta^2 \psi^{(2)} - \frac{\partial \psi^{(1)}}{\partial X} \\ & - 2\frac{\partial^3 \psi^{(1)}}{\partial t \partial x_p \partial X_p} + 4\nu\frac{\partial^2}{\partial x_p \partial X_p}\Delta_\perp \psi^{(1)} \\ & - \varepsilon_{jk}\frac{\partial \overline{\psi}}{\partial X_j}\frac{\partial}{\partial x_k}\Delta_\perp \psi^{(1)} = 0. \end{aligned} \tag{A6}$$

We calculate only the terms $\psi^{(0)}$, $\psi^{(1)}$, and $\psi^{(2)}$, since the terms $\psi^{(3)}$, $\psi^{(4)}$ do not contribute to Q, as will be seen below. Eqs. (A4), (A5) and (A6) can be solved using the Fourier transform over the fast variable x, that is, e.g.,

$$\psi^{(0)}(x,t)=\int d\mathbf{k}\hat{\psi}^{(0)}(k,t)e^{ikx}. \tag{A7}$$

Then, Eq. (A4) can be transformed to

$$\frac{\partial\hat{\psi}^{(0)}}{\partial t}+(i\omega_R+\nu_R)\hat{\psi}^{(0)}=\frac{F(\mathbf{k},t)}{1+k^2}, \tag{A8}$$

where

$$\omega_R=-\frac{k_x}{1+k^2},\quad \nu_R=\frac{\nu k^4}{1+k^2}.$$

The solution of Eq. (A8) is

$$\hat{\psi}^{(0)}(\mathbf{k},t)=\int_{-\infty}^{t}dt'\,\frac{F(\mathbf{k},t')}{1+k^2}\,\exp(i\Omega_k(t'-t)). \tag{A9}$$

Here $\Omega_k=\omega_R-i\nu_R$. The solutions of Eqs. (A5), (A6) can be obtained in a similar way.

$$\begin{aligned}\hat{\psi}^{(1)}(\mathbf{X},\mathbf{k},t)={}&-\varepsilon_{jk}\,\frac{ik_kk^2}{1+k^2}\,\frac{\partial\overline{\psi}}{\partial X_j}\int_{-\infty}^{t}dt'\\&\times\hat{\psi}^{(0)}(\mathbf{k},t')\exp(i\Omega_k(t'-t)),\end{aligned} \tag{A10}$$

$$\begin{aligned}\hat{\psi}^{(2)}(\mathbf{X},\mathbf{k},t)={}&-\varepsilon_{jk}\,\frac{ik_kk^2}{(1+k^2)^2}\,\frac{\partial^2\overline{\psi}}{\partial X\partial X_j}\\&\times\int_{-\infty}^{t}dt'e^{i\Omega_k(t'-t)}\int_{-\infty}^{t'}dt''\\&\times\hat{\psi}^{(0)}(\mathbf{k},t'')\exp(i\Omega_k(t''-t'))\\&+2\varepsilon_{jk}\,\frac{k_kk_pk^2}{(1+k^2)^2}\,\frac{\partial^2\overline{\psi}}{\partial X_p\partial X_j}\int_{-\infty}^{t}dt'\\&\times\hat{\psi}^{(0)}(\mathbf{k},t')\exp(i\Omega_k(t'-t))\\&-2\varepsilon_{jk}\,\frac{k_pk_kk^2}{(1+k^2)^2}\,i\Omega_k\,\frac{\partial^2\overline{\psi}}{\partial X_p\partial X_j}\int_{-\infty}^{t}dt'\\&\times\exp(i\Omega_k(t'-t))\int_{-\infty}^{t'}dt''\\&\times\hat{\psi}^{(0)}(\mathbf{k},t'')\exp(i\Omega_k(t''-t'))\\&+4\nu\varepsilon_{jk}\,\frac{k_pk_kk^4}{(1+k^2)^2}\,\frac{\partial^2\overline{\psi}}{\partial X_p\partial X_j}\int_{-\infty}^{t}dt'\\&\times\exp(i\Omega_k(t'-t))\int_{-\infty}^{t'}dt''\\&\times\hat{\psi}^{(0)}(\mathbf{k},t'')\exp(i\Omega_k(t''-t'))\\&-\varepsilon_{jk}\varepsilon_{lq}\,\frac{k_kk_qk^2}{(1+k^2)^2}\,\frac{\partial\overline{\psi}}{\partial X_l}\,\frac{\partial\overline{\psi}}{\partial X_j}\int_{-\infty}^{t}dt'\\&\times\exp(i\Omega_k(t'-t))\int_{-\infty}^{t'}dt''\\&\times\hat{\psi}^{(0)}(\mathbf{k},t'')\exp(i\Omega_k(t''-t')).\end{aligned} \tag{A11}$$

Now calculate the sum $Q^{(1)}=q^{(01)}+q^{(10)}$. Using the Fourier transform over the fast variables we get

$$\begin{aligned}q^{(01)}(\mathbf{X},t)={}&\varepsilon_{mn}\int\int d\mathbf{k}d\mathbf{k}'\,ik_m\Bigg[-ik_n'k'^2-2k_n'k_p'\,\frac{\partial}{\partial X_p}\\&+ik_n'\,\frac{\partial^2}{\partial X_p^2}-k'^2\,\frac{\partial}{\partial X_n}+2ik_p'\,\frac{\partial^2}{\partial X_p\partial X_n}\\&+\frac{\partial^3}{\partial X_p^2\partial X_n}\Bigg]\langle\hat{\psi}^{(0)}(\mathbf{k},t)\hat{\psi}^{(1)}(\mathbf{X},\mathbf{k}',t)\rangle\\&\times\exp(i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}).\end{aligned} \tag{A12}$$

It can be easily seen that some terms cancel due to

$$\varepsilon_{mn}k_mk_n=0.$$

Using Eq. (A10) we get

$$\begin{aligned}&\langle\hat{\psi}^{(0)}(\mathbf{k},t)\hat{\psi}^{(1)}(\mathbf{k}',\mathbf{X},t')\rangle\\&\quad=-\varepsilon_{jk}\,\frac{ik_k'k'^2}{1+k'^2}\,\frac{\partial\overline{\psi}}{\partial X_j}\int_{-\infty}^{t'}dt_1\langle\hat{\psi}^{(0)}(\mathbf{k},t)\\&\quad\times\hat{\psi}^{(0)}(\mathbf{k}',t_1)\rangle\exp(i\Omega_{k'}(t'-t_1)).\end{aligned} \tag{A13}$$

Here the assumption about the slow time evolution of $\overline{\psi}$ in comparison with the turbulent term $\hat{\psi}^{(0)}$ is used. This allows us to take the term $\partial\overline{\psi}/\partial X_j$ out of the integral.

For homogeneous turbulence we have

$$\langle\hat{\psi}^{(0)}(\mathbf{k},t)\hat{\psi}^{(0)}(\mathbf{k}',t_1)\rangle=\Phi(\mathbf{k},t-t_1)\delta(\mathbf{k}+\mathbf{k}'). \tag{A14}$$

Then we use the Fourier transform over time,

$$\Phi(\mathbf{k},t-t_1)=\int d\omega\Phi(\mathbf{k},\omega)\exp(-i\omega(t-t_1)). \tag{A15}$$

Inserting Eqs. (A15), (A14) into Eq. (A13) and then into Eq. (A12), we can calculate $q^{(10)}(\mathbf{X},t)$.

Analogous calculations are employed for $q^{(10)}(\mathbf{X},t)$, which is equal to

$$\begin{aligned}q^{(10)}(\mathbf{X},t)={}&\varepsilon_{mn}\int\int d\mathbf{k}d\mathbf{k}'\left(ik_m+\frac{\partial}{\partial X_m}\right)(-ik_n'k'^2)\\&\times\langle\hat{\psi}^{(1)}(\mathbf{X},\mathbf{k},t)\hat{\psi}^{(0)}(\mathbf{k}',t)\rangle\exp(i(\mathbf{k}+\mathbf{k}')\mathbf{x}).\end{aligned} \tag{A16}$$

Then, $Q^{(1)}$ has the form

$$\begin{aligned}Q^{(1)}=q^{(01)}+q^{(10)}={}&\varepsilon_{mn}\varepsilon_{jk}\Bigg(\eta_{mpk}\,\frac{\partial^3\overline{\psi}}{\partial X_p\partial X_n\partial X_j}\\&-\nu_{mk}^{(1)}\,\frac{\partial^2\Delta_s\overline{\psi}}{\partial X_n\partial X_j}\Bigg),\end{aligned} \tag{A17}$$

where η_{mpk}, $\nu_{mk}^{(1)}$ are determined according to Eqs. (2.13). For $Q^{(2)}$ one gets

$$Q^{(2)}=q^{(02)}+q^{(20)},$$

$$q^{(02)}=(\mathbf{X},t)=\varepsilon_{mn}\left\langle \frac{\partial\psi^{(0)}}{\partial x_m}\left(\frac{\partial}{\partial x_n}\,\Delta_\perp+2\,\frac{\partial^3}{\partial x_p\partial x_n\partial X_p}\right.\right.$$
$$\left.\left.+\frac{\partial}{\partial x_n}\,\Delta_s+\frac{\partial}{\partial X_n}\,\Delta_\perp+2\,\frac{\partial^3}{\partial x_p\partial X_p\partial X_n}\right)\psi^{(2)}\right\rangle, \tag{A18}$$

$$q^{(20)}(\mathbf{X},t)=\varepsilon_{mn}\left\langle\left(\frac{\partial}{\partial x_m}+\frac{\partial}{\partial X_m}\right)\psi^{(2)}\,\frac{\partial}{\partial x_n}\,\Delta_\perp\psi^{(0)}\right\rangle. \tag{A19}$$

Carrying out a Fourier transform over the fast variable and inserting the solution for $\hat{\psi}^{(2)}$ into Eqs. (A18) and (A19), we get

$$Q^{(2)}(\mathbf{X},t)=\varepsilon_{mn}\varepsilon_{jk}\left\{\nu^{(2)}_{mpk}\,\frac{\partial}{\partial X}\left(\frac{\partial^3\overline{\psi}}{\partial X_p\partial X_n\partial X_j}\right)\right.$$
$$-\nu^{(3)}_{mlpk}\,\frac{\partial^4\overline{\psi}}{\partial X_l\partial X_n\partial X_p\partial X_j}$$
$$\left.+\varepsilon_{rq}N^{(1)}_{mlkq}\,\frac{\partial^2}{\partial X_l\partial X_n}\left(\frac{\partial\overline{\psi}}{\partial X_r}\,\frac{\partial\overline{\psi}}{\partial X_j}\right)\right\}. \tag{A20}$$

It is easily seen that the sums $q^{(03)}+q^{(30)}$, $q^{(12)}+q^{(21)}$ and $q^{(04)}+q^{(40)}$ vanish by symmetry. Then the sum $q^{(13)}+q^{(31)}$ is equal to zero to fourth order in K. Therefore, only an expression for $q^{(11)}$ need be obtained:

$$q^{(11)}(X,t)=\varepsilon_{mn}\left\langle\left(\frac{\partial\psi^{(1)}}{\partial x_n}\,\frac{\partial}{\partial X_n}\,\Delta_\perp\psi^{(1)}\right.\right.$$
$$+2\,\frac{\partial\psi^{(1)}}{\partial x_m}\,\frac{\partial^3\psi^{(1)}}{\partial x_m\partial X_p\partial X_n}+\frac{\partial\psi^{(1)}}{\partial X_m}\,\frac{\partial}{\partial x_n}\,\Delta_\perp\psi^{(1)}$$
$$+2\,\frac{\partial\psi^{(1)}}{\partial X_m}\,\frac{\partial^3\psi^{(1)}}{\partial x_p\partial x_n\partial X_p}$$
$$\left.\left.+\frac{\partial\psi^{(1)}}{\partial X_m}\,\frac{\partial}{\partial X_n}\,\Delta_\perp\psi^{(1)}\right)\right\rangle. \tag{A21}$$

Inserting $\hat{\psi}^{(1)}$ (see Eq. (A10)) into Eq. (A21) we get

$$q^{(11)}(\mathbf{X},t)=\varepsilon_{mn}\varepsilon_{jk}\varepsilon_{rq}\left(N^{(2)}_{kpmq}\,\frac{\partial\overline{\psi}}{\partial X_j}\,\frac{\partial^3\overline{\psi}}{\partial X_p\partial X_n\partial X_r}\right.$$
$$\left.+N^{(3)}_{kpnq}\,\frac{\partial^2\overline{\psi}}{\partial X_m\partial X_j}\,\frac{\partial^2\overline{\psi}}{\partial X_p\partial X_r}\right). \tag{A22}$$

Now, summing up expressions for $Q^{(1)}$, $Q^{(2)}$ and $q^{(11)}$ we get the final expression for Q to order K^4; see Eq. (2.10).

APPENDIX B

Eddy viscosity of 2*D* NS flows: a comparison with previous results

In this Appendix we demonstrate the transition from the formula of Sec. 2 and Appendix A to those describing the turbulent viscosity of a 2D viscous incompressible fluid. It allows us, firstly, to compare our results with those of other authors studying turbulent viscosity in hydrodynamics in the framework of two-scale expansion, and, secondly, to point out the differences arising between calculations of the turbulent viscosity in 2D NS flows and in Rossby waves; see Sec. 4.

The coefficients in the equation for the mean stream function of viscous incompressible fluid are obtained from Eqs. (2.13) by setting $1+k^2\rightarrow k^2$ (no dispersion) and $\omega_R=0$ (zero eigenfrequency). We do not discuss effects nonlinear in ψ and, therefore, neglect all nonlinear terms. Further, we have $\nu^{(2)}_{mpk}=0$ because this term is determined by that with $\partial/\partial \underline{x}$ in Eq. (2.1). Therefore, we have the following equation for ψ

$$\frac{\partial}{\partial t}\,\Delta\overline{\psi}-\nu\Delta^2\overline{\psi}+\varepsilon_{mn}\varepsilon_{jk}\left(\eta_{mpk}\frac{\partial^3\overline{\psi}}{\partial X_p\partial X_n\partial X_j}\right.$$
$$\left.-\nu^{(1)}_{mk}\frac{\partial^2}{\partial X_n\partial X_j}\Delta_s\overline{\psi}+\nu^{(3)}_{mlpk}\frac{\partial^4\psi}{\partial X_l\partial X_n\partial X_p\partial X_j}\right)=0, \tag{B1}$$

where

$$\eta_{mpk}=\int d\mathbf{k}d\omega\frac{\Phi(\mathbf{k},\omega)}{\omega^2+\nu^2k^4}2\omega k_mk_pk_k,$$
$$\nu^{(1)}_{mk}=\int d\mathbf{k}d\omega\frac{\Phi(\mathbf{k},\omega)}{\omega^2+\nu^2k^4}\nu k^2k_mk_k, \tag{B2}$$
$$\nu^{(3)}_{mk}=\int d\mathbf{k}d\omega\frac{\Phi(\mathbf{k},\omega)}{(\omega^2+\nu^2k^4)^2}8\nu^3k^4k_mk_lk_pk_k.$$

Naturally, these results can be obtained if we start from Eq. (3.1) and use the method described in Appendix A.

Thus, the sign and the value of the turbulent viscosity are determined by the terms with $\nu^{(1)}_{mk}$, $\nu^{(3)}_{mlpk}$. As is shown in Sec. 3, the term with $\nu^{(1)}_{mk}$ is due to the transport of turbulent voriticity fluctuations by the mean flow. Even for the isotropic small-scale fluctuations this term makes a negative contribution to the turbulent viscosity. However, the terms with $\nu^{(3)}_{mlpk}$ can, of course, compensate the negative contribution.

At first, we note that the results which stem from Eqs. (B1) and (B2) are in accord with the results of Ref. 12. Indeed, let an external source in Eq. (3.1) be homogeneous in space and uncorrelated in time. Then, instead of Eqs. (B2) we have

$$\nu^{(1)}_{mk}=\int d\mathbf{k}\frac{k_mk_k\langle\hat{W}_0^2\rangle_k}{k^4\quad 2\nu k^2},$$
$$\nu^{(3)}_{mlpk}=\int d\mathbf{k}\frac{k_mk_lk_pk_k}{k^6}\,\frac{3\langle\hat{W}_0^2\rangle_k}{\nu k^2}, \tag{B3}$$

where $\langle W_0^2\rangle_k$ is the spatial spectrum of the zeroth-order approximation of the small-scale field vorticity. Equation (B1) leads to the equation for the Fourier transform of the large-scale vorticity $\widehat{\overline{W}}(\mathbf{K},t)$:

$$\frac{\partial}{\partial t}\,\widehat{\overline{W}}(\mathbf{K},t)=\gamma(\mathbf{K})\widehat{\overline{W}}(\mathbf{K},t)-\nu K^2\widehat{\overline{W}}(\mathbf{K},t), \tag{B4}$$

where

$$\gamma(\mathbf{K})=\int d\mathbf{k}\langle\hat{W}_0^2\rangle_k\,\frac{[\mathbf{K}\times\mathbf{k}]_z^2}{2\nu k^6}\left(1-\frac{6(\mathbf{K}\times\mathbf{k})^2}{k^2K^2}\right). \tag{B5}$$

Equation (B5) coincides with the main term of the expansion in powers of K/k of Eq. (2.15) of Ref. 12.

The case of isotropic small-scale turbulence is the simplest, so we start just from this one. We also use the Lorentzian shape for the spectral line broadening:

$$\Phi(\mathbf{k},\omega)=\frac{1}{\pi}\,\frac{\gamma_k}{\omega^2+\gamma_k^2}\,\Phi(k).$$

The Lorentzian shape is generally used: however, one may convince oneself that the result is not changed qualitatively by using another shape instead of Lorentzian one, for example, Gaussian.

$$\Phi(\mathbf{k},\omega)=\frac{1}{\sqrt{2\pi}\gamma_k}\exp\left(-\frac{\omega^2}{2\gamma_k^2}\right)\Phi(k), \tag{B6}$$

or the ‘‘step-like’’ shape,

$$\Phi(\mathbf{k},\omega)=\begin{cases}\dfrac{1}{2\gamma_k}\,\Phi(k), & |\omega|\leqslant\gamma_k\\ 0, & |\omega|\geqslant\gamma_k\end{cases}.$$

For the isotropic spectrum we have $\eta_{mpk}=0$. When calculating $\nu_{mk}^{(1)}$, $\nu_{mlpk}^{(3)}$ we use the following subsidiary integrals over the azimuthal angle φ of the wavenumber $\mathbf{k}$:

$$\int_0^{2\pi} d\varphi k_m k_n=\pi k^2\delta_{mn},$$

$$\int_0^{2\pi} d\varphi k_m k_l k_p k_k=\frac{\pi}{4}\,k^4(\delta_{ml}\delta_{pk}+\delta_{mp}\delta_{lk}+\delta_{mk}\delta_{lp}). \tag{B7}$$

Then

$$\nu_{mk}^{(1)}=\nu^{(1)}\delta_{mk},$$

$$\nu_{mk}^{(3)}=\nu^{(3)}(\delta_{ml}\delta_{pk}+\delta_{mp}\delta_{lk}+\delta_{mk}\delta_{lp}), \tag{B8}$$

where

$$\nu^{(1)}=\pi\int dk\;k^3\,\frac{\Phi(k)}{\nu k^2+\gamma_k},$$

$$\nu^{(3)}=\pi\int dk\;k^3\Phi(k)\,\frac{2\nu k^2+\gamma_k}{(\nu k^2+\gamma_k)^2}. \tag{B9}$$

Equation (B1) has the form

$$\frac{\partial}{\partial t}\,\Delta_s\overline{\psi}=\nu_{\text{eff}}\Delta_s^2\overline{\psi}, \tag{B10}$$

where

$$\nu_{\text{eff}}=\nu-\nu^{(1)}+\nu^{(3)}=\nu+\pi\int dk\;k^3\Phi(k)\,\frac{\nu k^2}{(\nu k^2+\gamma_k)^2}. \tag{B11}$$

It follows from the expressions obtained that the isotropic small-scale fluctuations do not give rise to a negative viscosity in the framework of the 2D NS equations for viscous incompressible fluid. The negative contribution $-\nu^{(1)}$ is compensated by the positive contribution $\nu^{(3)}$ arising in an accurate calculation of all the viscous terms in the framework of our scheme.

If we set $\gamma_k\to 0$ in Eq. (B11), then instead of Eq. (B10) we get

$$\frac{\partial}{\partial t}\,\Delta_s\overline{\psi}=\left(\frac{\langle\psi^2\rangle}{2\nu}+\nu\right)\Delta_s^2\overline{\psi}, \tag{B12}$$

where $\langle\psi^2\rangle$ is the variance of small-scale fluctuations. This result has been obtained by another method in Ref. 13.

[1] E. N. Lorenz, *The Nature and the Theory of the General Circulation of the Atmosphere*, WMO (1967).

[2] V. P. Starr, *Physics of Negative Viscosity Phenomena*, McGraw-Hill, New York (1968).

[3] R. H. Kraichnan, J. Atmos. Sci. **33**, 1521 (1976).

[4] A. Pouquet, J. Fluid Mech. **88**, 1 (1978).

[5] L. D. Meshalkin and Ya. B. Sinai, Prikl. Mat. Mekh. **25**, 1700 (1961).

[6] J. S. A. Green, J. Fluid Mech. **62**, 273 (1974).

[7] V. I. Yudovich, Prikl. Mat. Mekh. **29**, 453 (1965).

[8] U. Frisch, Z. S. She, and P. L. Sulem, Physica D **28**, 382 (1987).

[9] S. S. Moiseev, P. B. Rutkevich, A. V. Tur, and V. V. Yanovsky, Zh. Éksp. Teor. Fiz. **94**, 144 (1988) [Sov. Phys. JETP **67**, 513 (1988)].

[10] A. A. Nepomnyashchii, Prikl. Mat. Mekh. **40**, 886 (1976).

[11] G. I. Sivashinsky, Physica D **17**, 243 (1985).

[12] D. Montgomery and T. Hatori, Plasma Phys. Controlled Fusion **26**, 717 (1984).

[13] B. Dubrulle and U. Frisch, Phys. Rev. A **43**, 5355 (1991).

[14] A. V. Tur, A. V. Chechkin, and V. V. Yanovsky, Phys. Fluids B **4**, 3513 (1992).

[15] R. A. Madden, Rev. Geophys. Space Phys. **17**, 1935 (1979).

[16] A. R. Robinson, *Eddies in Marine Sciences*, Springer-Verlag, Berlin, Heidelberg, New York (1983).

[17] P. C. Liewer, Nucl. Fusion **25**, 543 (1985).

[18] J. La Belle, M. C. Kelley, and C. T. Surko, J. Geophys. Res. **91**, A5,5513 (1986).

[19] A. Hasegawa and C. Maclennan, Phys. Fluids **22**, 2212 (1979).

[20] J. Pedlosky, *Geophysical Fluid Dynamics*, Springer-Verlag, New York (1979).

[21] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed., Pergamon, Oxford (1982); Nauka, Moscow (1986).

[22] A. Hasegawa and K. Mima, Phys. Fluids **21**, 87 (1978).

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

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Combinatorial analysis of Feynman diagrams in problems with a Gaussian random field

É. Z. Kuchinskiĭ and M. V. Sadovskiĭ

Institute of Electrophysics, Ural Division, Russian Academy of Sciences, 620049 Ekaterinburg, Russia

(Submitted 16 May 1997)
Zh. Éksp. Teor. Fiz. **113**, 664–678 (February 1998)

We construct an algorithm for calculating the generating function for the number of skeleton graphs of the irreducible self-energy and vertex parts in the diagram technique for problems with a Gaussian random field. The exact recursion relation, defining the number of graphs in any order of perturbation theory, and the asymptotics in the high-order limit are found. The results obtained are applied to an analysis of the problem of an electron in a Gaussian random field with a white-noise correlator. A closed integral equation for the one-electron Green's function, the kernel of which is determined by the generating function, can be constructed in the approximation of equal skeleton graphs for the self-energy part in a given order of perturbation theory. An analysis shows that the approximation considered gives a qualitatively correct description of the tail of the state density in the region of negative energies and, probably, is fully applicable in the most interesting region of strong scattering near the edge of the original band where the asymptotics of the Green's function and the state density can be determined in the limit of infinitely strong scattering. © *1998 American Institute of Physics.* [S1063-7761(98)02102-7]

1. INTRODUCTION

Methods of summing Feynman diagrams are widely used in the consideration of a broad class of problems in theoretical physics in which the propagation of elementary perturbations (or quasiparticles) in statistically random fields created by an inhomogeneity is investigated. The simplest example of such a system is an electron propagating in a system of impurity atoms. It was precisely for this problem that the diagram technique to be considered in this paper was first formulated.[1,2] A similar technique is used in considering problems of statistical radiophysics and optics associated with the propagation of electromagnetic waves in disordered media.[3] The equivalent mathematical approach is applicable for a number of problems in the theory of critical phenomena in disordered systems,[4] in the problem of a polymer chain with an excluded volume and other problems in the physics of polymer systems.[5] Exactly the same diagram technique describes the regular model of critical phenomena with a zero-component order parameter.[4]

Information about the combinatorial analysis of graphs, i.e., about the number of diagrams of a given type in a given order of perturbation theory, is extremely useful in considering problems associated with the summation of Feynman diagrams. In this paper we will investigate in detail the question of the combinatorial analysis of diagrams in the above-mentioned class of problems.

2. GENERATING FUNCTION OF SKELETON DIAGRAMS: RECURSION RELATION

To be specific we will discuss the problem of an electron with energy E and momentum $\mathbf{p}$, propagating in a Gaussian random field (a system of random impurities).[1,2] The average one-particle Green's function is defined by the diagram series shown in Fig. 1a. This expansion is reduced in the usual fashion to the Dyson form:

$$G(E,p)=\frac{1}{E-\varepsilon_p-\Sigma(E,p)}, \tag{1}$$

where $\varepsilon_p=p^2/2m$ is the spectrum of a free electron, and the eigen-energy part $\Sigma(E,p)$ is determined by the skeleton graphs of Fig. 1b, in which the interior electron line represents the total (or dressed) Green's function $G(E,p)$.

The total number of graphs in the Nth order of perturbation theory in the expansion of Fig. 1a, as it is easy to see, is equal to

$$G_N=(2N-1)!!=\frac{(2N-1)!}{2^{N-1}(N-1)!}, \tag{2}$$

this is determined simply by the number of methods of connecting $2N$ vertices by N impurity lines. The problem of determining the analogous number of graphs Σ_N in the expansion of Fig. 1b is much more complicated, and as far as we know there is no exact answer in the literature. The simple inequality

$$(2N-1)!!>\Sigma_N>(2N-3)!!, \tag{3}$$

was found in Ref. 6, which only gives a fairly rough estimate of the quantity Σ_N. As we will see, the problem can be solved exactly. This follows directly from the exact solution of the problem of an electron in a random potential $V(\mathbf{r})=V$, where the quantity V does not depend on the spatial coordinate r but has a Gaussian distribution with width $\langle V^2\rangle=W^2$. It is natural that in this case the diagram technique has the standard form of Fig. 1, and each line of impurity interaction transfers zero momentum, i.e., it corre-

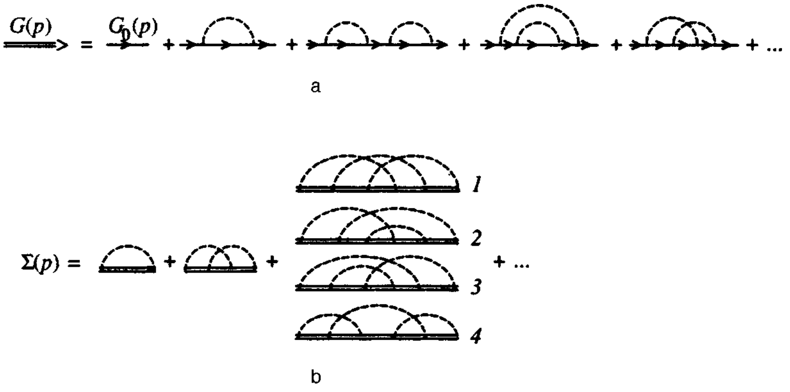


FIG. 1. Diagram series for average one-electron Green's function (a) and self-energy part (b). Dashed line corresponds to mean-square correlator of random field, G_0 is the free Green's function.

sponds (in the momentum representation) to the correlator $(2\pi)^d W^2 \delta(\mathbf{q})$ (d is the dimensionality of space).[7,8] All contributions of the same order in the expansion of Fig. 1a turn out to be the same, and the series for the Green's function is represented in the form[7]

$$G(E,p)=G_0(E,p)\left\{1+\sum_{N=1}^{\infty}(2N-1)!!G_0^{2N}(E,p)W^{2N}\right\}. \tag{4}$$

Then through the use of the representation

$$(2N-1)!!=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}dt\; t^{2N-2}e^{-t^2/2} \tag{5}$$

the series (4) is easily summed and we obtain[1)]

$$G(E,p)=\frac{1}{W}\,\Psi\left(\frac{1}{WG_0(E,p)}\right), \tag{6}$$

where the function

$$\Psi(z)=-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}dt\; e^{-t^2/2}\,\frac{1}{t-z}, \tag{7}$$

has been introduced.

Let us consider the self-energy part, corresponding to the Green's function (6). Since the addition of an impurity line leads in this problem simply to the additional multiplier W^2G^2, the self-energy part defined by the expansion of Fig. 1b can be written in the form

$$\Sigma=Q(W^2G^2)W^2G, \tag{8}$$

where $Q(x)$ is some function. We will see that this function is the generating function of the number of skeleton graphs for the self-energy part, i.e., its Taylor series expansion coefficients give the desired numbers Σ_N.

Let us write the Dyson equation for the problem being considered:

$$G=G_0+G_0\Sigma G=G_0(1+Q(W^2G^2)W^2G^2). \tag{9}$$

Introducing $z=(WG_0)^{-1}$ and $y=W^2G^2$, we obtain the following parametric representation of $Q(y)$ from Eqs. (6) and (9):

$$\begin{aligned}&1+yQ(y)=z\Psi(z)=z\sqrt{y},\\&y=\Psi^2(z).\end{aligned} \tag{10}$$

This representation of the function Q is rather inconvenient. Let us show that a differential equation can be obtained for it. It is easy to prove that the function $\Psi(z)$ satisfies the usual dispersion relation[2)]

$$\mathrm{Re}\,\Psi(z)=\frac{1}{\pi}\int_{-\infty}^{\infty}dt\,\frac{\mathrm{Im}\,\Psi(t)}{t-z},$$

$$\frac{1}{\pi}\,\mathrm{Im}\,\Psi(t)=\mp\frac{1}{\sqrt{2\pi}}\,e^{-t^2/2}, \tag{11}$$

from which it follows immediately that $\Psi(z)$ satisfies the differential equation

$$\frac{d\Psi}{dz}=1-z\Psi \tag{12}$$

with the initial condition

$$\Psi(z=\pm i0)=\mp i\sqrt{\pi/2}\,. \tag{13}$$

Differentiating the first equation in (10) with respect to y, we obtain

$$\frac{dz}{dy}=\frac{1}{2}\,y^{-3/2}\left\{2y^2\,\frac{dQ(y)}{dy}+yQ(y)-1\right\}. \tag{14}$$

Differentiating the second equation in (10) with respect to z and using Eq. (12), we have

$$\begin{aligned}\frac{dy}{dz}&=2\Psi(z)\,\frac{d\Psi(z)}{dz}=2\Psi(z)(1-z\Psi(z))\\&=-2y^{3/2}Q(y).\end{aligned} \tag{15}$$

By equating Eqs. (14) and (15), we obtain a nonlinear differential equation for $Q(y)$:

$$\frac{dQ(y)}{dy}=\frac{1}{2y^2}\,\{1-Q^{-1}(y)+y(Q)(y)\}. \tag{16}$$

Using Eqs. (10) and (13), we obtain $y=\Psi^2(z)|_{z=\pm i0}=-\pi/2$, so that

$$Q\left(-\frac{\pi}{2}\right)=\frac{z\Psi(z)-1}{7}\bigg|_{z=\pm i0}=\frac{2}{\pi}, \tag{17}$$

TABLE I.

N	$\Gamma_N=a_N$	$b_N=a_N/(2N+1)!!$	$\Sigma_N=a_{N-1}$	$U_N=(2N-1)a_{N-1}$
1	1	0.3333	1	1
2	4	0.2667	1	3
3	27	0.2571	4	20
4	248	0.2624	27	189
5	2830	0.2722	248	2232
6	38232	0.2829	2830	3120
7	593859	0.2930	38232	497016
8	10401712	0.3019	593859	8907885
9	202601898	0.3158	10401712	176829104
10	4342263000	0.3211	202601898	3849436062
$N\gg 1$	$\frac{1}{e}\left[1-\frac{5}{4N}\right](2N+1)!!$	$\frac{1}{e}\left[1-\frac{5}{4N}\right]$	$\frac{1}{e}\left[1-\frac{5}{4N}\right](2N-1)!!$	$\frac{1}{e}\left[1-\frac{9}{4N}\right](2N+1)!!$

which is the initial condition for Eq. (16). Note that the point $Q(0)=1$, with an obviousness that follows from the diagram representation for Σ, is a singular point for Eq. (16) and cannot serve as the initial condition.

Equation (16) can be rewritten in a form that is more convenient for further analysis

$$Q(y)=1+y\,\frac{d}{dy}\,yQ^2(y). \tag{18}$$

We are interested in the Taylor series expansion of $Q(y)$:

$$Q(y)=\sum_{n=0}^{\infty} a_n y^n. \tag{19}$$

Since the number of skeleton diagrams of Nth order for the self-energy part is simply the coefficient for W^{2N} in the series expansion of Σ in powers of W^2, it is easy to see that Eq. (8) gives the desired Σ_N in the form

$$\Sigma_N=a_{N-1}. \tag{20}$$

This also means that the function $Q(y)$ is the generating function for the combinatorial factors Σ_N of interest to us.

The substitution of Eq. (19) into (18) leads to the following recursion relation for the coefficients a_n:

$$a_n=n\sum_{m=0}^{n-1} a_m a_{n-1-m}, \tag{21}$$

where $a_0=1$. It follows directly from $a_0=1$ that $Q(0)=1$. It is precisely in this sense that this point is singular—the relation $Q(0)=1$ is satisfied for any initial conditions for which Eq. (18) has a solution.

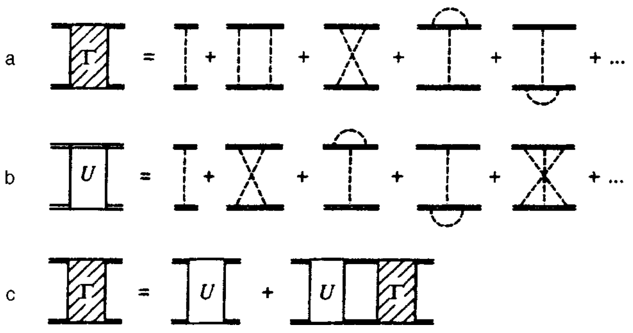


FIG. 2. Diagram series for the total vertex part Γ (a), for the irreducible vertex U (b), and the Bethe–Salpeter equation interrelating Γ and U (c).

From Eq. (21) it is easy to find the a_n values for small n; the corresponding results are listed in Table I.

By knowing the combinatorial analysis of the diagrams for the self-energy part, we can easily reproduce the combinatorial analysis for the two-particle Green's function—both for the total vertex part Γ and for the irreducible vertex U, the diagram expansion for which is given in Fig. 2. Actually, the self-energy part Σ is related to the total vertex Γ by the equation represented graphically in Fig. 3. For a problem with zero transferred momentum[7,8] this equation has the form

$$\Sigma=W^2G(1+G^2\Gamma). \tag{22}$$

Therefore, for the number of Nth-order diagrams in the total vertex Γ_N we obtain immediately

$$\Gamma_N=\Sigma_{N+1}=a_N. \tag{23}$$

Thus, the function $Q(y)$ is also the generating function for the number of diagrams of the total vertex part.

The number of Nth-order diagrams for the irreducible vertex U_N can easily be obtained if it is noted that a break of any of the $2N-1$ interior Green's lines in the diagram for the Nth-order self-energy part generates the corresponding diagram for the Nth-order contribution to the irreducible vertex U (Fig. 4). Therefore,

$$U_N=(2N-1)\Sigma_N=(2N-1)a_{N-1}. \tag{24}$$

In the Appendix we rederive the differential equation (18) for the generating function $Q(y)$ using only the Bethe–

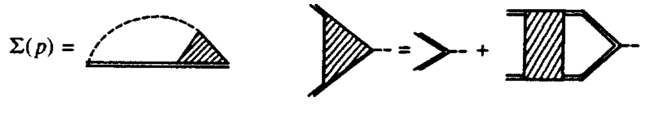


FIG. 3. Equations relating the eigen-energy part to the total vertex.

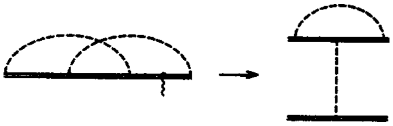


FIG. 4. Breaking of any of $2N-1$ interior lines of the Green's function in Nth-order skeleton diagram for self-energy part produces corresponding diagram for U.

Salpeter equation, which relates U to Γ, and the Ward identity, without using the explicit form of the Green's function (6).

3. ASYMPTOTICS FOR THE NUMBER OF DIAGRAMS FOR LARGE *N*

In the high-order limit $N \gg 1$ it becomes inconvenient to use the recursion relation (21) in view of the factorial increase in the number of diagrams.[6] At the same time the very fact of factorial growth can be used for a considerable simplification of the problem. We rewrite Eq. (21) in the form

$$a_n = n a_0 a_{n-1} + n a_1 a_{n-2} + n a_2 a_{n-3} + \dots, \tag{25}$$

where $a_0 = 1$, $a_1 = 1$, $a_2 = 4$. It is natural to assume that in the limit of large n we have $a_n \approx (2n+\beta) a_{n-1}$; then $a_{n-2} \approx a_{n-1}/(2n-2+\beta)$, etc. The substitution of these expressions into Eq. (25) immediately gives $\beta = 1$ and

$$a_n = \left(2n + 1 + O\left(\frac{1}{n}\right) \right) a_{n-1}. \tag{26}$$

This means that in the limit of large n we have $a_n \sim (2n+1)!!$ We define b_n as

$$b_n = \frac{a_n}{(2n+1)!!}. \tag{27}$$

Substituting Eq. (27) into (21), we obtain a recursion relation for b_n:

$$b_n = n \sum_{m=0}^{n-1} \frac{(2m+1)!!(2n-2m-1)!!}{(2n+1)!!} b_m b_{n-1-m}, \tag{28}$$

and $b_0 = 1$. In the limit of large n and taking into account $b_1 = 1/3$, $b_2 = 4/15$, which limits the accuracy to the order of b/n^2 (where $b \sim b_n \sim b_{n-1} \sim b_{n-2} \sim b_{n-3}$), we obtain

$$\Delta b_n = b_n - b_{n-1} = \frac{5}{4} \frac{b_{n-1}}{n^2} + O\left(\frac{b}{n^3}\right). \tag{29}$$

Thus, in the limit of large n we can write the following differential equation for b_n:

$$\frac{db_n}{dn} = \frac{5}{4} \frac{b_n}{n^2} + O\left(\frac{b}{n^3}\right), \tag{30}$$

from which it immediately follows that

$$b_n = b \exp\left(-\frac{5}{4}\frac{1}{n} + O\left(\frac{1}{n^2}\right) \right) = b\left\{ 1 - \frac{5}{4}\frac{1}{n} + O\left(\frac{1}{n^2}\right) \right\}. \tag{31}$$

Of course, on the basis of such an analysis it is impossible to determine the constant $b = \lim b_n$ as $n \to \infty$. A numerical analysis of the behavior of b_n using the recursion relation (28) completely corroborates the relationship (31) (see Fig. 5) and gives $b = 1/e = 0.36787944...$ (Calculations were carried out up to $n = 5000$, which ensures the stated accuracy.) We know of no analytical method for obtaining this curious result.

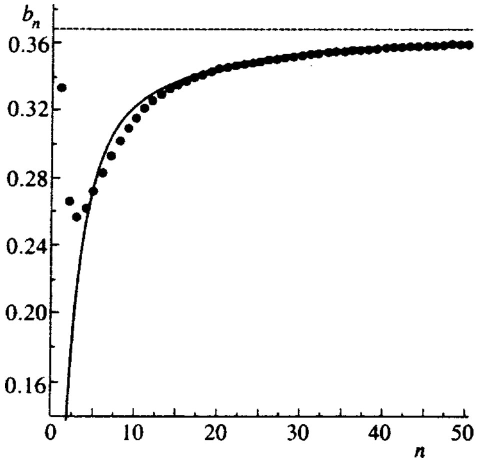


FIG. 5. Behavior of b_n with increase in n. Points correspond to b_n values obtained from recursion relation (28), the curve corresponds to the asymptotic dependence $e^{-1}(1-5/4n)$, the dashed line corresponds to the asymptotic function $1/e$.

Finally, the asymptotics of the number of diagrams of different types for large N have the form[3)]

$$\begin{aligned} \Sigma_N = a_{N-1} &= b_{N-1}(2N-1)!! \\ &= \frac{1}{e}\left\{ 1 - \frac{5}{4}\frac{1}{N} + O\left(\frac{1}{N^2}\right) \right\}(2N-1)!! \\ &= \frac{1}{\sqrt{\pi} e}\left\{ 1 - \frac{5}{4}\frac{1}{N} + O\left(\frac{1}{N^2}\right) \right\} 2^N \Gamma\left(N + \frac{1}{2}\right), \end{aligned} \tag{32}$$

$$\begin{aligned} \Gamma_N = a_N &= \frac{1}{e}\left\{ 1 - \frac{5}{4}\frac{1}{N} + O\left(\frac{1}{N^2}\right) \right\}(2N+1)!! \\ &= \frac{1}{\sqrt{\pi} e}\left\{ 1 - \frac{5}{4}\frac{1}{N} + O\left(\frac{1}{N^2}\right) \right\} 2^{N+1} \Gamma\left(N + \frac{3}{2}\right), \end{aligned} \tag{33}$$

$$\begin{aligned} U_N = (2N-1)a_{N-1} &= \frac{1}{e}\left\{ 1 - \frac{5}{4}\frac{1}{N} + O\left(\frac{1}{N^2}\right) \right\}(2N-1) \\ &\times (2N-1)!! = \frac{1}{e}\left\{ 1 - \frac{9}{4}\frac{1}{N} + O\left(\frac{1}{N^2}\right) \right\}(2N+1)!! \\ &= \frac{1}{\sqrt{\pi} e}\left\{ 1 - \frac{9}{4}\frac{1}{N} + O\left(\frac{1}{N^2}\right) \right\} 2^{N+1} \Gamma\left(N + \frac{3}{2}\right). \end{aligned} \tag{34}$$

It is interesting to note that

$$\frac{\Sigma_N}{G_N} = b_{N-1} = \frac{1}{e}\left\{ 1 - \frac{5}{4}\frac{1}{N} + O\left(\frac{1}{N^2}\right) \right\} \to \frac{1}{e}, \tag{35}$$

$$\frac{U_N}{\Gamma_N} = 1 - \frac{1}{N} + O\left(\frac{1}{N^2}\right) \to 1. \tag{36}$$

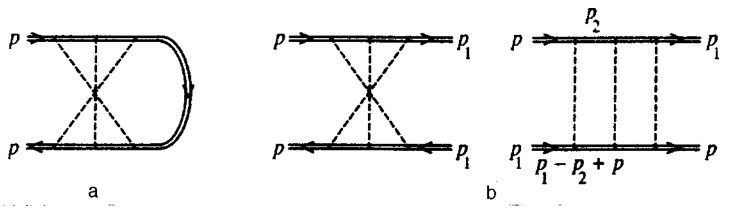


FIG. 6. (a)—base graph used in constructing approximation for eigen-energy part, (b)—expanded sequence of maximally intersecting graphs gives ladder in the case of system invariance with respect to time reversal operation.

Table I summarizes the principal results for the number of graphs of different types.

4. ELECTRON IN GAUSSIAN RANDOM FIELD WITH WHITE-NOISE CORRELATOR

As an example of the practical use of the results obtained above let us consider the problem of an electron in a Gaussian random field with a white-noise correlator when the impurity interaction line corresponds to the expression[1,2,9]

$$\omega(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4)=W^2\delta(\mathbf{p}_1-\mathbf{p}_2+\mathbf{p}_3-\mathbf{p}_4), \qquad (37)$$

where $W^2=\rho V^2$, ρ is the density of impurity atoms, and V is the Born amplitude of scattering at a point impurity. It is well-known that the principal difficulties in this problem arise at energies defined by the condition[9]

$$|E|\lesssim\gamma(E) \quad \text{or} \quad |E|\lesssim E_{sc}, \qquad (38)$$

where $\gamma(E)=\pi\rho V^2N(E)$ is the Born damping ($N(E)$ is the state density, corresponding to the energy E), $E_{sc}\sim m^{d/(4-d)}(\rho V^2)^{2/(4-d)}$ is the characteristic size of the critical region near the band edge, where strong scattering arises. These difficulties are associated primarily with the impossibility of selecting a particular dominant sequence of Feynman diagrams similarly to what is done in the weak scattering region[4)] $E\gg\gamma(E)$, $E\gg E_{\rm sc}$.[1,2] Actually, all diagrams for the self-energy part are of the same order in the $|E|\lesssim E_{\rm sc}$ region and must be taken into account.

The perturbation theory series for the self-energy part is shown in Fig. 1b in terms of skeleton graphs. By means of simple variable replacements one can show that all third-order graphs in this expansion are equal to one another (diagrams of Fig. 1b (1–4)). Although this equality breaks down in even the next order, it is reasonable to formulate an approximation in which it is assumed that all graphs of this type are equal in each order of perturbation theory. Such an approximation should give satisfactory results primarily in the critical region $|E|\lesssim E_{sc}$, where all contributions have at least the same order of magnitude. We choose as the base graph in each order the maximally intersecting type shown in Fig. 6a. The sequence of interaction lines entering into it for systems that are invariant with respect to time reversal can be transformed into a ladder, as shown in Fig. 6b. Then the complete series for the self-energy part in our approximation is represented in the form

$$\Sigma(p)=\sum_{n=1}^{\infty}W^2\Sigma_n\sum_{\mathbf{p}_1}\sum_{\mathbf{p}_2}[W^2G(\mathbf{p}_1+\mathbf{p}_2+\mathbf{p})G(-\mathbf{p}_2)]^{n-1}G(\mathbf{p}_1)$$

$$=\sum_{\mathbf{p}_1}W^2Q\left[W^2\sum_{\mathbf{p}_2}G(\mathbf{p}_1-\mathbf{p}_2+\mathbf{p})G(\mathbf{p}_2)\right]G(\mathbf{p}_1), \qquad (39)$$

where the definitions (19) and (20) were used, as well as the property $G(\mathbf{p})=G(-\mathbf{p})$ in an isotropic system. Correspondingly, we obtain the closed equation for the average one-particle Green's function in the form

$$G^{-1}(p)=G_0^{-1}(p)-W^2\sum_{\mathbf{q}}Q\left[W^2\sum_{\mathbf{p}_1}G(\mathbf{p}_1-\mathbf{q})\times G(\mathbf{p}_1)\right]G(\mathbf{p}+\mathbf{q}), \qquad (40)$$

where $G_0^{-1}(p)=E-p^2/2m$. The entire nontrivial part of the problem being considered is now expressed by means of the generating function $Q(y)$, which determines the kernel of the complex nonlinear integral equation (40). Restricting consideration to the first term of the expansion (19) gives $Q=1$, and Eq. (40) reduces to the standard problem of summing nonintersecting graphs.[1,2] An obvious advantage of the result (40) compared with the standard approach,[1,2] based on identifying the dominant sequence of diagrams (for example, taking account of only the first graph in Fig. 1b), is that it formally accounts for all diagrams, which is done, however, in the approximation that all skeleton graphs for the self-energy part are equal in a given order of perturbation theory.

Equation (40) is an extremely complicated nonlinear integral equation and cannot be solved in general form, and what is more we do not know the general form of the function $Q(y)$ (which, moreover, enters into Eq. (40) as a function of a complex argument). We will restrict ourselves below to some qualitative analysis of the consequences arising from Eq. (40). We write Eq. (40) in compact form as

$$G^{-1}(p)=G_0^{-1}(p)-W^2Q[W^2G\otimes G]\otimes G, \qquad (41)$$

where the generalized product (or convolution) of functions

$$F\otimes\Phi=\sum_{\mathbf{p}}F(\mathbf{p}-\mathbf{q})\Phi(\mathbf{p}), \qquad (42)$$

has been introduced, and we return to the system of Eqs. (10) which define the function Q parametrically. The second equation in (10) is now written as

$$G\otimes G=\frac{1}{W^2}\Psi^2(z). \qquad (43)$$

We saw above that $z=W^{-1}G_0^{-1}$ holds in the problem with zero transferred momentum. Let us examine the limit $W\to 0$ in Eq. (43). Then the left side of Eq. (43) is reduced to

$G_0\otimes G_0$, and on the right side one can assume, by analogy with the problem with zero transferred momentum, $z\sim W^{-1}$ and can use the asymptotic form $\Psi(z)\approx 1/z$ for $|z|\gg 1$. There is some error here since the exact form of $\Psi(z)$ is

$$\Psi(z)=R(z)\mp i\sqrt{\frac{\pi}{2}}e^{-z^2/2}, \tag{44}$$

where an asymptotic expansion of the form

$$R(z)=e^{-z^2/2}\int_0^z e^{t^2/2}dt=\frac{1}{z}+\frac{1}{z^3}+\frac{3}{z^5}+\ldots\left(-\frac{\pi}{4}<\arg z<\frac{\pi}{4}\right). \tag{45}$$

exists for $R(z)$. We use the asymptotic form $\Psi(z)\approx 1/z$, which is not completely true, but the results obtained by using this approximation are corroborated in a more rigorous but much more lengthy analysis. Thus, in the limit $W\to 0$ Eq. (43) reduces to

$$G_0\otimes G_0=\frac{1}{W^2z^2} \quad \text{or} \quad z=\frac{1+O(W^2)}{W\sqrt{G_0\otimes G_0}}. \tag{46}$$

Correspondingly, in the limit $W\to 0$ we can write

$$G\otimes G=\frac{1}{W^2}\,\Psi^2\left(\frac{1}{W\sqrt{G_0\otimes G_0}}\right). \tag{47}$$

in place of (43). Let us consider the energy region $E<0$, where the fluctuational tail of the state density arises.[9,10] In this case we have $z\in\mathrm{Re}$ from Eq. (46). By means of Eqs. (44) and (46) we obtain

$$G\otimes G\approx G_0\otimes G_0-i\,\frac{2}{W}\,\sqrt{\frac{\pi}{2}}\sqrt{G_0\otimes G_0}\times\exp\left\{-\frac{1}{2W^2G_0\otimes G_0}\right\}, \tag{48}$$

from (47), where, as we now see, the second term also produces a fluctuational tail of the state density. Using

$$\sum_{\mathbf{q}}G\otimes G=\sum_{\mathbf{p}}\sum_{\mathbf{q}}G(\mathbf{p}-\mathbf{q})G(\mathbf{p})=\left(\sum_{\mathbf{p}}G(p)\right)^2.$$

we obtain immediately from Eq. (48) the state density in the form

$$N(E)=-\frac{1}{\pi}\sum_{\mathbf{p}}\ \mathrm{Im}\ G^R(E,p)=\frac{1}{\sqrt{2\pi}W}\,\frac{\Sigma_{\mathbf{q}}\sqrt{G_0\otimes G_0}\ \exp\{-1/(2W^2G_0\otimes G_0)\}}{|\Sigma_{\mathbf{p}}G_0(E,p)|}. \tag{49}$$

Thereafter everything is determined by the specific form of $G_0\otimes G_0$ in spaces of different dimension.

In the one-dimensional $(d=1)$ case all of the integrals entering into Eq. (49) are calculated exactly. After rather involved but fairly elementary calculations we obtain

$$N(E)=\frac{1}{2\pi}\,\sqrt{\frac{2m}{|E|}}\,\exp\left\{-\sqrt{2}\,\frac{|E|^{3/2}}{m^{1/2}W^2}\right\}. \tag{50}$$

The argument of the exponential in Eq. (50) differs from the known exact result of Halperin[11] (see also Chapter 11 in Ref. 10) by the absence of a 4/3 multiplier. The pre-exponential function in Eq. (50) also differs from the exact, which is $\sim|E|/W^2$.[11] Nevertheless, the behavior of the state density tail is reproduced quite satisfactorily in a qualitative sense in our approximation. In this regard let us recall the widespread notion that the state density tail cannot be obtained at all from perturbation theory.

Analogous (but still approximate) calculations of the state density using Eq. (49) for $d=3$ yield

$$N(E)\sim\exp\left\{-\sqrt{2}\,\frac{|E|^{1/2}}{m^{3/2}W^2}\right\}. \tag{51}$$

Here the exponential once more coincides with the known result of the nonperturbative instanton approach within the accuracy of a constant.[9,12–14] The pre-exponential multiplier, omitted in Eq. (51), following from Eq. (49), does not coincide with any of the known versions obtained in the cited papers. Nevertheless, the result (51) for the dominant exponent is also quite satisfactory despite the approximate character of Eq.5) (40).

An analysis of the consequences of Eq. (40) in the strong-coupling region,[9] defined by condition (38), i.e., in the vicinity of the edge of the initial band where a transition from spatial to localized states occurs, is of special interest. There is every basis for assuming that in this region the approximation of equal contributions to the self-energy part in a given order of perturbation theory can turn out to be good simply because of the known fact that they are equal in order of magnitude. A strong condition of the type (38), obviously, is equivalent to passing to the limit $W\to\infty$. In this limit in the zeroth approximation one can ignore in Eq. (41) the first term on the right side compared with the second and can write

$$G^{-1}(p)=-W^2Q[W^2G\otimes G]\otimes G. \tag{52}$$

We see that this corresponds to the limit $z=\pm i0$ in Eq. (43) for $y=-\pi/2$ in Eq. (10). In this case Eq. (43) is reduced to

$$W^2G\otimes G=\Psi(z=\pm i0)=-\pi/2, \tag{53}$$

and we have from Eq. (17)

$$Q[W^2G\otimes G]=2/\pi. \tag{54}$$

The formal solution of Eq. (53) has the form

$$G=\pm i\sqrt{\frac{\pi}{2}}\frac{1}{W\sqrt{\mathcal{N}}}, \tag{55}$$

where $\mathcal{N}=\Sigma_{\mathbf{p}}1$ is the number of states in the band. It is easy to see that this equation is satisfied by the direct substitution of Eqs. (55) and (54) into (52). Thus, in a first approximation in the limit $W\to\infty$ one can write the Green's function (41) in the form

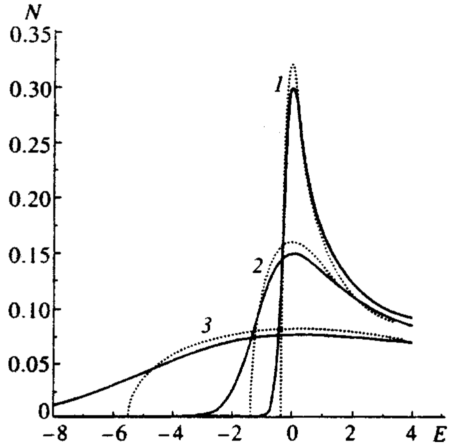


FIG. 7. State density in one-dimensional system for different values of the mean square of random field $W^2(2m)^{1/2}/E_0^{3/2}$: *1*—0.25, *2*—2, *3*—16. Solid curves represent exact solution, dotted curves represent self-consistent Born approximation (56). Energy is given in units of E_0 and the state density in units of $\sqrt{2m/E_0}$, where E_0 is arbitrary.

$$G(p)=\frac{1}{G_0^{-1}(p)-(2/\pi)W^2\Sigma_{\mathbf{p}}G(p)}, \tag{56}$$

which agrees surprisingly well with the result of the self-consistent Born approximation (the first diagram in Fig. 1b or Fig. 3)[1,2] to within the redundant multiplier $2/\pi$. Equation (56) leads in an obvious manner to the state density of the Born approximation $N_0(E)$, which practically coincides for $d=3$ with the state density of the free electron model (with a one-loop displacement of the band edge taken into account). Figure 7 shows a comparison of the results following from Eq. (56) for the state density in a one-dimensional ($d=1$) system with the exact Halperin result,[11] demonstrating satisfactory agreement of these results in the strong-coupling region $|E|<E_{sc}\sim m^{1/3}W^{4/3}$, the width of which increases with an increase in W. It must be pointed out that although the tail of the state density is suppressed with an increase in W (see Eq. (50)), the intermediate region where $|E|\sim E_{sc}$ increases.

It is possible that a result of the form (56) makes it possible to justify qualitatively using the simplest Born approximation for the one-electron Green's function in approaches such as the self-consistent localization theory[9,15]—the mobility threshold occurs in the strong-coupling region $|E|\lesssim E_{sc}$ (38), where the approximation (56) turns out to be quite satisfactory and the Green's function actually has the simple Born form.

This work was partially supported by the Russian Fund for Fundamental Research (Project 96-02-16065) and was also carried out within the framework of Project IX.I of the Statistical Physics Government Program of the Russian Ministry of Science. The authors are grateful to A. I. Posazhennikova for assisting with the numerical calculations.

APPENDIX

Let us derive Eq. (18) for the generating function $Q(y)$ without using the explicit form of the one-particle Green's function (6). In problems with zero momentum transfer momentum the Bethe–Salpeter equation of Fig. 2c has the form

$$\Gamma=U+UG^2\Gamma, \tag{A1}$$

so that

$$\Gamma=\frac{U}{1-UG^2}. \tag{A2}$$

Using Eqs. (A2) and (22), we obtain an equation relating the self-energy part to the irreducible vertex U:

$$\sum=\frac{W^2G}{1-UG^2}. \tag{A3}$$

We use the Ward identity

$$W^2\frac{\partial}{\partial W}\bigg|_G\frac{\Sigma}{W}=UG, \tag{A4}$$

the validity of which is easy to see by means of Eqs. (8) and (24), and Eq. (A2) in order to write

$$W^2\frac{\partial}{\partial W}\bigg|_G\frac{\Sigma}{W}=UG=\frac{1}{G}\left\{1-W^2\frac{G}{\Sigma}\right\}$$

or

$$\Sigma=W^2G+W^2G\Sigma\frac{\partial}{\partial W}\bigg|_G\frac{\Sigma}{W}. \tag{A5}$$

Using Eq. (8), we obtain the desired differential equation for Q:

$$\begin{aligned}Q(W^2G^2)&=1+W^2GQ(W^2G^2)\frac{\partial}{\partial W}\bigg|_G WGQ(W^2G^2)\\&=1+W^2G^2\frac{d}{d(W^2G^2)}W^2G^2Q^2(W^2G^2),\end{aligned}$$

which is rewritten as

$$Q(y)=1+y\frac{d}{dy}yQ^2(y). \tag{A6}$$

Note, however, that from these arguments it is impossible to find the correct boundary condition (17), which is closely related to the relation (11), reflecting the causality principle.

[1)] From a mathematical viewpoint this means Borel summation.

[2)] The sign of the imaginary part corresponds to treating the retarded or advanced Green's functions.

[3)] An asymptotic limit of the form (32), $\Sigma_N\approx c\cdot 2^N\Gamma(N+\beta)$, was obtained in Ref. 6 by the Lipatov method; however, the coefficients c and β were not found.

[4)] In this case the nonintersecting diagrams dominate, so that one can take account of only the first diagram in Fig. 1b.

[5)] For $d>4$ knowledge of the asymptotic form (32) and the statistical analysis of Ref. 6 make it possible to determine the correct exponent of W^{-1} in the pre-exponential function of the state density. In this case our approximation is equivalent to the hypothesis, used in Ref. 6, that the high-order contributions are stationary, which is valid for $d>4$.

---

[1] S. F. Edwards, Philos. Mag. **3**, 1020 (1958).

[2] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Physics* (Fizmatgiz, Moscow, 1962; Prentice-Hall, Englewood Cliffs, NJ, 1963), Chap 7.

[3] S. N. Rytov, Yu. A. Kravtsov, and V. N. Tatarskiĭ, *Principles of Statistical Radiophysics* (Nauka, Moscow, 1978, Springer, New York, 1987), Pt. II, Chap. VIII.

[4] S. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, Massachusetts, 1976; Mir, Moscow, 1980), Chap. 10.

[5] A. Yu. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (Nauka, Moscow, 1989; AIP, New York, 1994), Chap. 2.

[6] I. M. Syslov, Zh. Éksp. Teor. Fiz. **102**, 1951 (1992) [Sov. Phys. JETP **75**, 1049 (1992)].

[7] L. V. Keldysh, Dissertation...doctor phys. math. sciences (FIAN, Moscow, 1965).

[8] A. L. Efros and B. I. Shklovskiĭ, *Electronic Properties of Doped Semiconductors* (Springer, New York, 1984), Chap. 11.

[9] M. V. Sadovskii, Sov. Sci. Rev. A—Phys. Rev. **7**, 1 (1986).

[10] I. M. Lifshits, S. A. Gredeskul, and L. A. Pastur, *Introduction to the Theory of Disordered Systems* (Wiley, New York, 1988), Chap. IV.

[11] B. I. Halperin, Phys. Rev. **139**, A104 (1965).

[12] J. Cardy, J. Phys. C **11**, L321 (1978).

[13] M. V. Sadovskiĭ, Fiz. Tverd. Tela **21**, 743 (1979) [Sov. Phys. Solid State **21**, 435 (1979)].

[14] I. M. Suslov, Zh. Éksp. Teor. Fiz. **111**, 1896 (1997) [JETP **84**, 1036 (1997)].

[15] D. Vollhardt and P. Wölfe, in *Electronic Phase Transition*, ed. by W. Hanke and Yu. V. Kopaev (North-Holland, Amsterdam, 1992), p. 1.

Translated by Eugene R. Heath

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Vacancy formation energy in icosahedral structures

V. V. Kovshevnyĭ,*) D. V. Olenev, and Yu. Kh. Vekilov

Moscow State Institute of Steel and Alloys, 117936 Moscow, Russia

(Submitted 18 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 679–687 (February 1998)

A method is developed for calculating the formation energy of vacancies that is applicable to crystalline, amorphous, and quasicrystalline structures. Calculations are performed in the strong-binding approximation for Amman–Mackey networks with different types of decorations. It is shown that the most closely packed structure, for which the formation energy of vacancies is much less than the crystalline state, has the smallest vacancy formation energy. This can explain the experimental data on the increased concentration of vacancies in quasicrystals compared with crystalline objects [R. Chidambaram, M. K. Sanyal, V. S. Raghunathan, P. M. G. Nambissan, and P. Pen, Phys. Rev. B **48**, 3030 (1993)]. © *1998 American Institute of Physics.* [S1063-7761(98)02202-1]

1. INTRODUCTION

It has been found experimentally by the positron annihilation method that the concentration of point defects in quasicrystals is several orders of magnitude higher than in the crystalline case. The assumption was put forth that structural models of icosahedral quasicrystals based on the Amman-Mackey network (two-fragment model of the structural skeleton of a quasicrystal) cannot explain such concentrations of point defects.[1] However, this assumption was not corroborated by reliable physical arguments, primarily because of the topological complexity of the structure of icosahedral quasicrystals. The purpose of this paper is to show that an increased concentration of vacancies can be explained in terms of the two-fragment structural model of a quasiperiodic object. To do this a method is developed in this paper for calculating the formation energy of vacancies in icosahedral structures using a recursion (or continued fraction) method. Since the numerical calculations for an actual quasicrystal are rather lengthy, an Amman–Mackey network cluster, obtained by the projection of a six-dimensional space and decorated by iron atoms in different positions, was chosen as the object of the investigations in order to study the general relationships. The choice of iron made it possible to restrict consideration to the interaction of nearest neighbors in calculations by the strong-binding method, in view of the localized character of the *d* orbitals. The calculations were carried out for solitary vacancies (Schottky defects).

The layout of this paper is as follows. The calculation method is presented in Sec. 2. The results of calculations and their analysis are presented in Sec. 3. Section 4 contains conclusions.

2. CALCULATION METHOD

The formation energy of a vacancy is defined in the following manner:

$$E_V = E(N-1,V) - E(N-1,0), \tag{1}$$

where $E(N-1,V)$ and $E(N-1,0)$ are the energies of clusters of $N-1$ atoms with and without a vacancy, respectively. This expression can be rewritten in the form[2]

$$E_V = [E(N-1,V) - E(N,0)] + [E(N,0) - E(N-1,0)], \tag{2}$$

where $E(N,0)$ is the energy of a cluster of N atoms.

The first term inside the brackets corresponds to the energy needed to remove an atom from the cluster to infinity, the second term describes the change in the energy of an ideal cluster when the number of atoms is changed from $N-1$ to N. For large N one can assume

$$E(N,0) - E(N-1,0) \approx \frac{E(N,0)}{N}. \tag{3}$$

Substituting Eq. (3) into (2), we obtain

$$E_V = E(N-1,V) - E(N,0) + \frac{E(N,0)}{N}. \tag{4}$$

The total energy of a cluster can be written in the form of a sum, containing the attraction, caused by the covalent hybridization of the electron states at the different atoms, in addition to the Born–Meyer repulsion of ionic shells, determined in the pair approximation[3]:

$$E_{\text{tot}} = \frac{1}{2} \sum_{i,j}{}' \varphi(|r_{ij}|) + 2 \int_{-\infty}^{E_F} \rho(E) E\, dE. \tag{5}$$

The prime on the summation sign in the first term means that the term with $i=j$ is omitted; E_F is the Fermi energy. Denoting the first term in Eq. (5) by W, we can rewrite Eq. (4) in the following manner:

$$E_V = (W_1 - W_2) + 2\left\{ \int_{-\infty}^{E_{F_1}} \rho_1(E) E\, dE - \int_{-\infty}^{E_{F_2}} \rho_2(E) E\, dE \right\} + \frac{1}{N} \times \left(W_2 + 2 \int_{-\infty}^{E_{F_2}} \rho_2(E) E\, dE \right). \tag{6}$$

where the indices 1 and 2 refer, respectively, to clusters with and without a defect.

The strong-binding approximation, which is valid in the case of the transition metals, can be used for constructing the cluster Hamiltonian, which is required to calculate the second and third terms in Eq. (6). The Hamiltonian of the problem is

$$H=\sum_{i,\lambda} |i\lambda\rangle \mathscr{E}_d^{\,i} \langle i\lambda| + \sum_{\substack{i,j\\ \lambda,\mu}}{}' \; |i\lambda\rangle h_{i\lambda,j\mu}\langle j\mu|, \tag{7}$$

where i,j are indices labeling the atoms of the cluster; λ, μ are the indices of the atomic d orbitals; $\langle i\lambda|j\mu\rangle=\delta_{ij}\delta_{\lambda\mu}$ (basis orthogonality); and $h_{i\lambda,j\mu}$ are the Koster–Slater matrix elements in the two-center approximation.[4] The prime on the summation sign in the second term means that the term with $i=j$ is omitted. In view of the one-component nature of the system all atomic levels with energies $\mathscr{E}_d^{\,i}$ can be assumed to be identical and set equal to zero.

It is easy to show (see Appendix) that Eq. (6) with (7) taken into account can be rewritten as

$$\begin{aligned} E_V=&(W_1-W_2)+\frac{1}{N}\left\{W_2+2\int_{-\infty}^{E_F}\rho(E)E\,dE\right\}\\ &+2\sum_{\alpha=1}^{5}\left\{\sum_i z_{i\alpha}^{(2)}-\sum_i p_{i\alpha}^{(2)}+(N_{p\alpha}^{(2)}-N_{z\alpha}^{(2)})E_F\right\}, \end{aligned} \tag{8}$$

where ρ and E_F are the state density and Fermi energy of a defect-free cluster, calculated by means of a recursion method (see Appendix); $z_{i\alpha}^{(2)}$ (or $p_{i\alpha}^{(2)}$) are the zeros (or poles) of the matrix element of the resolvent, obtained from the Hamiltonian with $\alpha-1$ removed rows and columns, lying below the Fermi energy of the defect-free structure; and $N_{z\alpha}^{(2)}$ (or $N_{p\alpha}^{(2)}$) is the number of zeros $z_{i\alpha}^{(2)}$ (or poles $p_{i\alpha}^{(2)}$).

Relaxation effects in the formation of a vacancy were ignored in this paper. This approximation is justified to a certain degree for close-packed structures, where relaxation effects make a contribution to the vacancy formation energy amounting to a few percent.[5]

3. FORMATION ENERGY OF VACANCIES IN QUASICRYSTALLINE CLUSTERS

The formation energy of vacancies was calculated in the strong-binding approximation for three types of decoration of the Amman–Mackey network rhombohedrons by iron atoms. The Amman–Mackey network (three-dimensional Penrose packing), which is one of the most widely used models of the structural skeleton of quasicrystals, is obtained by means of two structural blocks,[6–9] acute and obtuse rhombohedrons. If these rhombohedrons are filled with atoms (atoms are placed at the vertices, at the centers of the faces, etc.), then the so-called decorated Amman–Mackey network is obtained. All of the packing is characterized by the edge length a_R of the rhombohedron, which for real quasicrystals is of the order of several angstroms (5–6 Å).[8,9]

The topology of the structure was investigated for each type of decoration—from 20 to 121 different local configurations of atoms were selected. The formation energy of a vacancy in BCC iron was calculated to verify the method. A cluster consisting of 1024 atoms was used. In order to suppress surface effects the calculations were performed for the interior portion of the cluster (the first three coordination spheres from the center site for BCC iron). Since iron is a $3d$ transition metal, the atoms of the first two coordination spheres were taken into account in the calculation of the electron contribution (increasing the size of the cluster and the number of coordination spheres taken into account affected the result only slightly). Although relaxation was ignored, it should be pointed out that the relaxation correction in BCC iron can amount to 40–50% of the calculated vacancy formation energy.[10] The calculated vacancy formation energy amounted to 1.94 eV, which agrees well with the experimental data, which lie in the energy interval from 1.3 eV to 1.6 eV for iron with the BCC structure.[11] Thus, the calculation results confirmed the good applicability of the method for normal crystalline objects.

For each cluster investigated (a fragment of an Amman–Mackey network was used as the cluster) the topology was analyzed beforehand in order to identify nonequivalent local atomic configurations.

The Henley classification[12] was used in the topological analysis of the primitive Amman–Mackey network (the atoms are located only at the vertices of the rhombohedrons) to denote the different local topological configurations. In the consideration of the atoms on the nearest three coordination spheres the local configuration is characterized by four indices $(\alpha\beta\gamma)_p$, where α is the number of a bonds coming out from a site (a is the radius of the first coordination sphere); β is the number of b bonds, coming out from a site (b is the radius of the second coordination sphere); γ is the number of c bonds, coming out from a site (c is the radius of the third coordination sphere); and p is the number of b bonds forming a nearly complete ring about the a bonds.[12]

According to this classification 27 different configurations can be identified in the primitive decoration case.

The frequency at which a local configuration is encountered can be determined from the formula

$$\xi_{(\alpha\beta\gamma)_p}=\lim_{N\to\infty}\frac{N_{(\alpha\beta\gamma)_p}}{N}, \tag{9}$$

where $N_{(\alpha\beta\gamma)_p}$ is the number of local configurations in the portion of the cluster being investigated, having a certain $(\alpha\beta\gamma)_p$; N is the number of atoms in the portion of the cluster being investigated.

In the case of more complex decorations the search for topologically nonequivalent local configurations of atoms was conducted in the following manner. Atoms located on the first three coordination spheres were also considered as nearest neighbors. Each local configuration of atoms was characterized by a set of three parameters: n_a, n_b and n_c, where n_a, is the number of atoms belonging to the first coordination sphere; n_b and n_c are the number of atoms on the second and third sphere, respectively (here the radii of the coordination spheres are different from the quantities a, b and c for the primitive decoration). The atom configurations were considered to be different if at least one of the

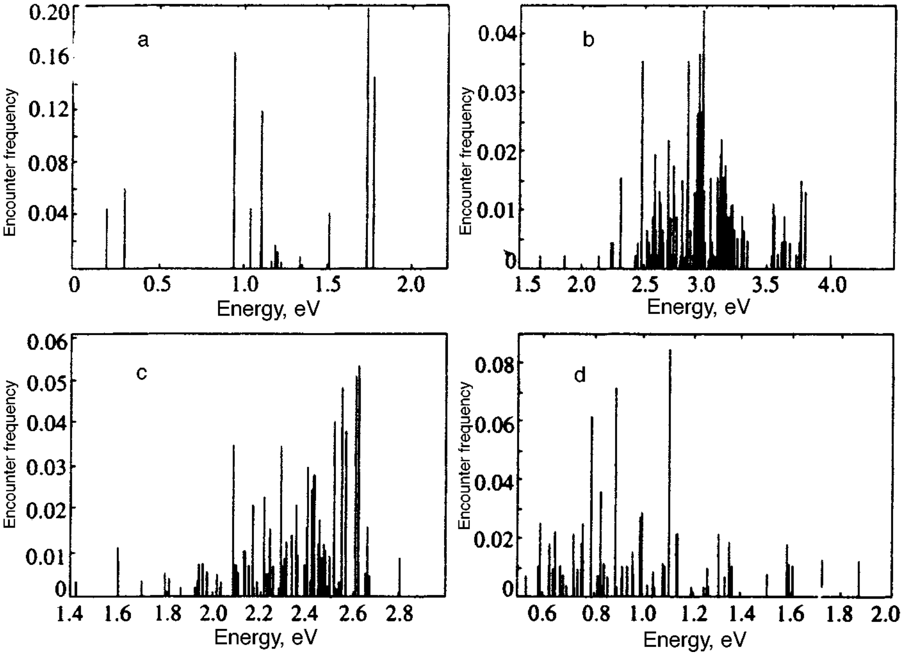


FIG. 1. Vacancy formation energy in decorated Amman–Mackey networks: a) primitive decoration (18 local nonequivalent configurations); b) atoms at centers of rhombohedron edges (121 local nonequivalent configurations); c) bonds leading to structure porosity of local nonequivalent configurations are ignored; d) atoms at vertices and at centers of rhombohedron edges (74 local nonequivalent configurations).

parameters did not coincide when n_a, n_b and n_c were compared for these configurations. It was found that in the central portion of the resulting cluster of 978 iron atoms, located at the centers of the rhombohedron edges, 121 nonequivalent configurations are realized.

In the case of the Amman–Mackey network, in which the bonds are lengthened, leading to structure porosity (c are the bonds mentioned above),[12] the topological analysis revealed 84 nonequivalent configurations of atoms in the central portion of the cluster.

For the decoration of the vertices and centers of the rhombohedron edges with atoms the procedure for identifying nonequivalent local configurations was altered in the following manner in view of the greater complexity of the structure. Each local configuration was characterized by five parameters: n_a, n_b, n_c, n_d and n_e, where n_a is the number of nearest neighbors lying on the first sphere, n_b on the second sphere, n_c, n_d, n_e are the number of atoms on the third, fourth and fifth spheres, respectively. Any two local atomic configurations were considered to be equivalent in the case when all five parameters n_a, n_b, n_c, n_d and n_e were equal. With this approach 74 nonequivalent local configurations of atoms are found in the investigated portion of the cluster with atoms lying at the vertices and centers of the rhombohedron edges.

Figure 1a–d shows the results of a calculation of the vacancy formation energy in clusters on the basis of the Amman–Mackey network with four different decorations. The large diversity of local configurations in the decorated Amman–Mackey networks leads to a wide range of vacancy formation energies. In the primitive Amman–Mackey network (Fig. 1a) the vacancy formation energy varies from 0.2 to 1.8 eV, and the configurations with a small vacancy formation energy have a high encounter frequency. In an Amman–Mackey network with iron atoms at the centers of the rhombohedron edges (Fig. 1b) most of the vacancy formation energies lie in the interval of values from 2.4 to 3.3 eV, although some energies lie near a value of 4 eV. Obviously, taking relaxation corrections into account should reduce these values significantly because of the fact that this structure has a low structure packing factor. In the case of an Amman–Mackey network with excluded minimal bonds the results show (Fig. 1c) that the spread in vacancy formation energy is reduced. Most of the vacancy formation energies lie within the interval from 2.1 to 2.7 eV, i.e., the local configurations have become more alike. When atoms are located at the vertices and at the centers of the rhombohedron edges (Fig. 1d), the spread in the vacancy formation energies spans the interval from 0.49 eV to 1.83 eV. This interval has become narrower than in the two previous cases; this indicates an increased filling homogeneity of the space in the cluster and a reduction in the contribution of relaxation corrections.

Thus, one can conclude that an increased concentration of vacancies in quasicrystals can be explained by the existence of a large fraction of local configurations in the decorated Amman–Mackey networks that have a rather low vacancy formation energy.

4. CONCLUSIONS

The method we have developed for calculating the formation energy of vacancies has made it possible to investigate the effect of the type of decoration of the Amman–Mackey network on the spectrum of vacancy formation energies. An investigation of the vacancy formation energy was performed for four types of decoration of the Amman–Mackey network with only iron atoms in order to determine the general relationships. A preliminary calculation for crystalline BCC iron showed excellent agreement with experiment. The small vacancy formation energies in a quasicrystal compared with the crystalline case arise because the decorated Amman–Mackey network has a large set of nonequivalent local topological configurations, in which the vacancy formation energy varies within wide limits. With an increase in the degree of packing and the homogeneity of the quasiperiodic structure the interval of vacancy formation energies is shifted to lower energies; this is explained by a reduction in the role of the relaxation corrections. Thus, the results of our analysis show that the experimentally observed increase in the concentration of point defects in quasicrystals is explainable within the framework of the model of a decorated Amman–Mackey network with ''favorable'' sites existing in it for the formation of vacancies. Calculations for specific quasicrystalline alloys are the subject of a separate paper.

APPENDIX

Equation (6) for the vacancy formation energy is transformed into the final Eq. (8) in the following manner.

We denote the expression inside the braces in Eq. (6) by $\{\}$.

Since the model to be considered for the formation of a defect in a quasicrystalline cluster presumes conservation of the number of electrons, we have

$$\int_{-\infty}^{E_{F_1}} \rho_1(E)\,dE = \int_{-\infty}^{E_{F_2}} \rho_2(E)\,dE. \tag{A1}$$

With Eq. (A1) taken into account, we rewrite $\{\}$ in the following manner:[13]

$$\{\} = \int_{-\infty}^{E_F} (E-E_F)(\rho_1(E)-\rho_2(E))\,dE. \tag{A2}$$

The index on E_F is omitted, since the small displacement of the Fermi level between the two structures has already been taken into account in the derivation of Eq. (A2). Now the quantity E_F means the Fermi energy of an ideal cluster. In a contour-integral representation Eq. (A2) is written in the form

$$\{\} = \frac{1}{2\pi i} \oint_c (E-E_F)(\mathrm{Tr}[EI-H_1]^{-1} - \mathrm{Tr}[EI-H_2]^{-1})\,dE, \tag{A3}$$

where c is the integration contour in the complex plane, encompassing all poles $\mathrm{Tr}[EI-H_2]^{-1}$ lying below the Fermi energy E_F of the ideal cluster; H_1 and H_2 are the Hamiltonians of the cluster with and without defects.

We denote

$$N = \mathrm{Tr}[EI-H]^{-1}. \tag{A4}$$

Then Eq. (A3) is rewritten as

$$\{\} = \frac{1}{2\pi i} \oint_c (E-E_F)\Delta N\,dE, \tag{A5}$$

where $\Delta N = N_1 - N_2$.

Using the well-known relation

$$\mathrm{Tr}[EI-H]^{-1} = \frac{\partial}{\partial E} \ln(\det[EI-H]), \tag{A6}$$

we can write

$$\Delta N = N_1 - N_2 = \frac{\partial}{\partial E} \ln \frac{\det[EI-H_1]}{\det[EI-H_2]}. \tag{A7}$$

Let us examine the block structure of the matrices of the Hamiltonians H_1 and H_2:

$$H_1 = H_0,$$

$$H_2 = \begin{bmatrix} E_A & \tilde{V}_1 \\ V_1 & H_0 \end{bmatrix}. \tag{A8}$$

Here E_A is the matrix of resonance integrals for an isolated atom A; V_1 is the matrix describing the interaction of the atom A being considered with the rest of the cluster; and H_0 is the matrix describing the rest of the cluster. We can write[13]

$$\det\,[EI-H_1] = \det\,[EI-H_0],$$

$$\det\,[EI-H_2] = \det\,G_2^{-1}\,\det\,[EI-H_0], \tag{A9}$$

where G_2 is an $n \times n$ submatrix in the upper left corner of the matrix $G=[EI-H_2]^{-1}$ (when only d orbitals are taken into account we have $n=5$). The expression for $\det G_2$ can be written in the form[13]

$$\det\,G_2 = \overline{G}_0 \overline{G}_1 \overline{G}_2 \ldots \overline{G}_{n-1}, \tag{A10}$$

where $\overline{G}_i = [EI-H_{2i}]_{11}^{-1}$, i.e., this is the element in the upper left corner of the matrix $[EI-H_{2i}]^{-1}$, H_{2i} is an $(N-i)\times(N-i)$ matrix, formed by removing the first i rows and columns of the matrix H_2, $i=0, n-1$ $(n=5)$.

Then

$$\Delta N = \frac{\partial}{\partial E} \ln(\det\,G_2). \tag{A11}$$

If

$$(G_2)_{\alpha\beta} = \langle \alpha | [EI-H_2]^{-1} | \beta \rangle, \tag{A12}$$

where α and β are basis d orbitals, then, taking Eq. (A10) into account, we write

$$\det\,G_2 = \overline{G}_{21} \overline{G}_{22} \overline{G}_{23} \overline{G}_{24} \overline{G}_{25}, \tag{A13}$$

where $\overline{G}_{2\alpha} = \langle \alpha | [EI-\overline{H}_{2\alpha}]^{-1} | \alpha \rangle$, $\alpha=1,\ldots,5$. Here $\overline{H}_{2\alpha}$ is the matrix formed by removing the first $(\alpha-1)$ rows and columns in H_2. Substituting Eq. (A13) into (A11), we have

$$\Delta N = \sum_{\alpha=1}^{5} \frac{\partial}{\partial E} \ln \overline{G}_{2\alpha}. \tag{A14}$$

A recursion method[14] can be used to determine $\overline{G}_{2\alpha}$ $(\alpha=1\ldots5)$. If $|\varphi\rangle$ is a localized orbital, then the matrix element of the resolvent operator $G=\langle \varphi | [EI-H]^{-1} | \varphi \rangle$ can be represented in the form of the continued fraction

$$G = \cfrac{1}{E - a_0 - \cfrac{b_1}{E - a_1 - \cfrac{b_2}{\cdots}}} \tag{A15}$$

We write the continued fraction, calculated up to the coefficients a_p, b_{p+1}, in the form

$$G=\frac{(p-1)\text{th degree polynomial}}{\rho\text{th degree polynomial}} \tag{A16}$$

or

$$G=\frac{\Pi_{i=1}^{p-1}(E-z_i)}{\Pi_{i=1}^{p}(E-p_i)}, \tag{A17}$$

where z_i (or p_i) are the zeros (or poles) of G. Then Eq. (A14) is rewritten as

$$\Delta N=\sum_{\alpha=1}^{5}\left\{\sum_{i=1}^{p-1}\frac{1}{E-z_{z\alpha}^{(2)}}-\sum_{i=1}^{p}\frac{1}{E-p_{i\alpha}^{(2)}}\right\}, \tag{A18}$$

where $z_{i\alpha}^{(2)}$ (or $p_{i\alpha}^{(2)}$) are the zeros (or poles) of the matrix element of the resolvent $G_{2\alpha}$.

Substituting Eq. (A18) into (A5), we have

$$\{\ \}=\sum_{\alpha=1}^{5}\left\{\sum_{i} z_{i\alpha}^{(2)}-\sum_{i} p_{i\alpha}^{(2)}+(N_{p\alpha}^{(2)}-N_{z\alpha}^{(2)})E_F\right\}, \tag{A19}$$

where $N_{p\alpha}^{(2)}$ (or $N_{z\alpha}^{(2)}$) are the number of poles (or zeros) of the matrix element of the resolvent of the ideal cluster, obtained from the matrix of the Hamiltonian with $\alpha-1$ removed rows and columns. The zeros and poles lying below the Fermi energy of the defect-free cluster are taken into account.

As a result, we obtain the final formula (8).

*)E-mail: vict28@yahoo.com

---

[1] R. Chidambaram, M. K. Sanyal, V. S. Raghunathan, P. M. G. Nambissan, and P. Pen, Phys. Rev. B **48**, 3030 (1993).

[2] M. Sinder, D. Fuks, and J. Pelleg, Phys. Rev. B **50**, 2775 (1994).

[3] W. Eckstein, *Computer Modeling of the Interaction of Particles with a Solid Surface* (Mir, Moscow, 1995).

[4] J. S. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

[5] M. J. Mehl and B. M. Klein, Physica B **172**, 211 (1991).

[6] D. V. Olenev, P. A. Korzhavy, and Yu. Kh. Vekilov, Zh. Éksp. Teor. Fiz. **104**, 4130 (1993) [JETP **77**, 998 (1993)].

[7] V. Elser, Acta Crystallogr., Sect. A: Found. Crystallogr. **42**, 36 (1986).

[8] V. Elser, Phys. Rev. B **32**, 4892 (1985).

[9] D. Levine and P. J. Steinhardt, Phys. Rev. B **34**, 596 (1986).

[10] V. G. Vaks and A. V. Trefilov, Fiz. Tverd. Tela **19**, 244 (1977) [Sov. Phys. Solid State **19**, 139 (1977)].

[11] M. A. Shtremel', *Strength of Alloys, Part 1, Lattice Defects* (Metallurgiya, Moscow, 1973).

[12] C. L. Henley, Phys. Rev. B **34**, 797 (1986).

[13] N. R. Burke, Surf. Sci. **58**, 349 (1976).

[14] R. Haidock, V. Heine, and M. Kelly, J. Phys. 8, 2591 (1975).

Translated by Eugene R. Heath

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Transverse runaway of hot electrons and the electron-temperature approximation

Z. S. Kachlishvili and F. G. Chumburidze

Iv. Dzhavakhishvili Tbilisi State University, 380028 Tbilisi, Georgia

(Submitted 30 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 688–692 (February 1998)

This paper discusses the transverse runaway effect in the electron-temperature approximation. The combinations of scattering mechanisms and the corresponding threshold electric fields for which transverse runaway develops are determined. It is shown that the transverse-runaway effect is not associated with any approximation. © *1998 American Institute of Physics.* [S1063-7761(98)02302-6]

Let us consider a semiconductor in crossed strong electric and magnetic fields **E** and **H**. The isotropic part of the nonequilibrium distribution function of hot electrons for quasi-elastic energy scattering by various types of phonons and for momentum scattering by phonons or defects can be written as (see, for example, Ref. 1)

$$f_0(x)\propto\exp\left[-\frac{dx}{1+E^2Q(x)}\right],\tag{1}$$

where

$$Q(x)=\frac{1}{E_0^2}\,\frac{x^{(t+s)/2}}{1+\eta x^t}\tag{2}$$

is the heating function,

$$E_0\equiv\frac{\sqrt{3}k_0T}{e(l_0\tilde{l}_0)^{1/2}},\quad \eta=\left(\frac{H}{H_0}\right)^2,\quad H_0\equiv\frac{(2mc^2k_0T)^{1/2}}{el_0},$$

$x=\varepsilon/k_0T$, k_0T is the thermal energy, and e and m are the charge and effective mass of the electron. It is assumed that the energy dependence of the mean free paths with respect to momentum, l, and with respect to energy, $\overline{l}$, can be written as

$$l(x)=l_0x^{(1+t)/2},\quad \tilde{l}=\tilde{l}_0x^{(1+s)/2},\tag{1a}$$

The values of t and s for all known scattering mechanisms are given in Ref. 1.

As can be seen from Eqs. (1) and (2), the nonequilibrium distribution function depends on the scattering mechanisms via the heating function and the internal (applied plus Hall) field. In his study of the asymptotic behavior of the heating function, Levinson[2] classified the types of runaway. However, it was shown in Refs. 3 and 4 that there is a new type of runaway of hot electrons, associated with the dependence of the internal (heating) field on the scattering mechanisms. This type of runaway was named *transverse runaway*.[3] It has a threshold character. Combinations of energy- and momentum-scattering mechanisms have been found for which transverse runaway occurs with threshold values both in applied electric fields and in magnetic fields.

The question naturally arises whether the transverse-runaway effect has a universal character or whether it is associated with the hot-electron quasi-elastic scattering approximation. To find the answer to this question, we have studied runaway in the electron-temperature approximation. This article presents the results of these studies.

Following the electron-temperature method, we assume that the nonequilibrium scattering function is the Maxwellian function

$$f(x)=A\,\exp\left(-\frac{x}{\Theta}\right),\tag{3}$$

where $\Theta=T_e/T$, T_e is the electron temperature, and A is a normalizing factor. We shall determine the electron temperature from the energy-balance equation[5]

$$\mathbf{j}\cdot\mathbf{E}=nk_0T\,\frac{\Theta-1}{\langle\tau_e\rangle},\tag{4}$$

where n is the electron density, $\langle\tau_e\rangle$ is the mean energy-relaxation time, $\mathbf{j}$ is the current density:

$$\mathbf{j}=-en\left\{\mu_1\mathbf{E}+\mu_2\,\frac{\mathbf{E}\times\mathbf{H}}{H}\right\}.\tag{5}$$

The mobilities μ_1 and μ_2 are determined in terms of $f(x)$:

$$\frac{\mu_1}{\mu_0}=\frac{\Gamma(3/2)}{\Gamma((t+5)/2)}\,\frac{J_1}{J_0},\quad \frac{\mu_2}{\mu_0}=\frac{\Gamma(3/2)}{\Gamma((t+5)/2)}\,\sqrt{\eta}\,\frac{J_2}{J_0},\tag{6}$$

where

$$J_1=\int_0^\infty\left(-\frac{\partial f}{\partial x}\right)\frac{x^{(t+3)/2}}{1+\eta x^t}\,dx,\tag{7}$$

$$J_2=\int_0^\infty\left(-\frac{\partial f}{\partial x}\right)\frac{x^{(2t+3)/2}}{1+\eta x^t}\,dx,\tag{8}$$

$$J_0=\int_0^\infty x^{1/2}f\,dx,\tag{9}$$

μ_0 is the mobility in zero electric field, and $\Gamma(t)$ is the gamma function.

Using Eq. (1a), for $\langle\tau_e\rangle$ we get

$$\langle\tau_e\rangle=\tau_e^0\,\frac{\Gamma((3+s)/2)}{\Gamma(3/2)}\,\Theta^{s/2},\tag{10}$$

where $\tau_e^0=\overline{l}_0(2mk_0T)^{1/2}$.

Let there be a rectangular semiconductor sample and let an electric field E_x be applied and a current j_x flow along the x axis. A magnetic field is also applied along the z axis. We consider the regime in which the current direction is given (open Hall contacts). Then the Hall field E_y is determined from the condition $j_y=0$, while the energy-balance equation takes the form

$$E_x^2\,\frac{\Gamma((3+s)/2)}{\Gamma(3/2)}\,\frac{e\mu_0\tau_e^0}{k_0T}\,\frac{J_1}{J_0}\left[1+\eta\left(\frac{J_2}{J_1}\right)^2\right]\Theta^{s/2}-\Theta+1=0. \tag{11}$$

It can be seen from this that the electron temperature Θ is a function of the applied electric and magnetic fields and the parameters t and s. There is interest in the question: Is there a combination of scattering mechanisms t and s for which the solution of Eq. (11) goes to infinity as a function of one of the parameters E_x, H? In the limit $\Theta\to\infty$, the distribution function is not normalized; i.e., runaway of the hot electrons occurs.

An analytical solution of Eq. (11) is possible only in the approximation of strong $(h\overline{x^t}\gg 1)$ and weak $(h\overline{x^t}\ll 1)$ magnetic fields.

Strong magnetic field case. Dropping the 1 in the denominators of Eqs. (7) and (8) and computing the integrals J_1 and J_2, we can rewrite the energy-balance equation as

$$a_1E_x^2\Theta^{(t+s)/2}+\frac{a_2}{\eta}\,E_x^2\Theta^{(s-t)/2}-\Theta+1=0, \tag{12}$$

where

$$a_1\equiv\frac{e\mu_0\tau_e^0}{k_0T}\,\frac{\Gamma^2(5/2)}{\Gamma(3/2)\Gamma((s-t)/2)},$$

$$a_2\equiv\frac{e\mu_0\tau_e^0}{k_0T}\,\frac{\Gamma((s-t)/2)}{\Gamma(3/2)}.$$

The condition for the appearance of runaway with respect to the applied electric field, $\partial\Theta/\partial E_x\to\infty$ (with $H=\mathrm{const}$), is fulfilled when

$$a_1\,\frac{t+s}{2}\,E_x^2\Theta^{(t+s-2)/2}+\frac{a_2}{\eta}\,\frac{s-t}{2}\,E_x^2\Theta^{(s-t-2)/2}=1. \tag{13}$$

Consequently, Eq. (13) is the equation for the asymptotic values of Θ. When it is satisfied for a definite E_x value, we have $\Theta\to\infty$. We first find the combinations of scattering mechanisms for which we have $\Theta\to\infty$ for finite E_x. To do this, we separately consider positive and negative values of t for any values of s.

(1) Let $t>0$. Then the inequality $t+s-2>-t+s-2$ is always satisfied. In this case, the first term in Eq. (13) is the main one. Dropping the second term and determining E_x^2, it is easy to convince oneself that, when $\Theta\to\infty$, the finiteness of E_x is ensured only for scattering mechanisms that satisfy the condition $t+s=2$.

Substituting $t>0$ and $t+s=2$ into Eq. (12), we find for the asymptotic solution of this equation

$$\Theta=\frac{1}{1-(E_x/E_{01})^2}. \tag{14}$$

Runaway takes place in the limit $E_x\to E_{01}$; $E_{01}\ldots E_{04}\sim(k_0T/e\mu_0\tau_e^0)^{1/2}$ is the characteristic electric field.

(2) Now let $t<0$. It is obvious that $s-|t|-2<s+|t|-2$, and the main term in this case is the second term, while the finiteness of E_x as $\Theta\to\infty$ is ensured by the scattering mechanisms $s+|t|=2$.

Taking $t<0$ and $s+|t|=2$ into account in Eq. (12), we get for the asymptotic solution of Θ

$$\Theta=\left[1-\frac{1}{\eta}\left(\frac{E_x}{E_{02}}\right)^2\right]^{-1}. \tag{15}$$

Runaway occurs as $E_x\to\sqrt{\eta}E_{02}$.

Thus, in a strong magnetic field, runaway develops for the scattering mechanisms

$$t>0,\quad t+s=2;\quad t<0,\quad s+|t|=2.$$

The threshold value of E_x here is independent of the magnetic field in the first case, and it depends on the magnetic field in the second case.

Weak magnetic field case. Because of the smallness of ηx^t in comparison with unity, we drop it out of the denominators of Eqs. (7) and (8), after which the energy-balance equation takes the form

$$b_1\eta E_x^2\Theta^{(3t+s)/2}+b_2E_x^2\Theta^{(s+t)/2}-\Theta+1=0, \tag{16}$$

where

$$b_1\equiv\frac{e\mu_0\tau_e^0}{k_0T}\,\frac{\Gamma((3+s)/2)}{\Gamma(3/2)}\,\frac{\Gamma^2((2t+5)/2)}{\Gamma^2((t+5)/2)},$$

$$b_2=b_1\,\frac{\Gamma^2((t+s)/2)}{\Gamma^2(2t+s/2)}.$$

The runaway condition $\partial\Theta/\partial E_x\to\infty$ is given by

$$b_1\eta\,\frac{3t+s}{2}\,\Theta^{(3t+s-2)/2}+b_2E_x^2\,\frac{s+t}{2}\,\Theta^{(s+t-2)/2}=1. \tag{17}$$

Carrying out a treatment analogous to the strong magnetic field case, we find the following:

(1) For $t>0$, the finiteness of the applied electric field is ensured by the condition $3t+s=2$, and, from Eq. (16), we get for the asymptotic solution

$$\Theta=\frac{1}{1-\eta(E_x/E_{03})^2}. \tag{18}$$

Runaway occurs as $E_x\to E_{03}/\sqrt{\eta},$.

(2) In the case $t<0$, runaway arises for $s-|t|=2$, while the asymptotic value of Θ has the form

$$\Theta=\frac{1}{1-(E_x/E_{04})^2}. \tag{19}$$

Runaway occurs as $E_x\to E_{04}$.

Thus, runaway develops under strong magnetic field conditions when

$$t>0,\quad t+s=2;\quad t<0,\quad s+|t|=2;$$

and it develops under weak magnetic field conditions when

$$t>0,\quad 3t+s=2;\quad t<0,\quad s-|t|=2.$$

Let us now clarify the character of these runaways. As already said, transverse runaway is the name applied to runaway for which the heating field grows without limit because of the Hall field. The heating field can be written as

$$E^2=E_x^2[1+(\mu_2/\mu_1)^2].$$

Computing μ_1 and μ_2 in the strong and weak magnetic field approximations, we get, respectively,

$$E^2=E_x^2\left[1+\frac{\Gamma^2(5/2)}{\Gamma^2((5-t)/2)}\ \eta\Theta^t\right]$$

and

$$E^2=E_x^2\left[1+\frac{\Gamma^2((5+t)/2)}{\Gamma^2((5+2t)/2)}\ \eta\Theta^t\right].$$

It can be seen from these expressions that transverse runaway takes place only for $t>0$ and $t+s=2$, $3t+s=2$ for the threshold values of the applied electric field given by Eqs. (14) and (18), respectively, and this precisely coincides with the threshold values of E_x obtained in the quasi-elastic scattering approximation.[3,4] As far as negative t is concerned, transverse runaway does not develop for these values, since we have $E\rightarrow E_x$ for runaway [$\Theta\rightarrow\infty$ in Eqs. (15) and (19)]; i.e., no Hall field is formed in this case. It is easy to convince oneself by investigating the asymptotic forms of the heating functions given by Eq. (2) that runaway takes place for negative t, associated with the so-called limited-restraining scattering mechanism.[2]

We conclude from this that transverse runaway substantially differs from other types of runaway. It is associated with unlimited growth of the heating (Hall) field. This is physically associated with the fact that, for $t>0$, as the mean energy increases, the scattering frequency goes to zero and the Lorentz force acts more effectively.

It can apparently be concluded from these results that transverse runaway is a universal effect. It is not associated with any approximation and occurs for the same combinations of scattering mechanisms and for the same threshold values of the applied fields.

We are deeply grateful to I. P. Zvyagin and A. G. Mironov for discussing the results and for a number of useful comments.

[1] Z. S. Kachlishvili, Phys. Status Solidi A **33**, 15 (1976).
[2] I. B. Levinson, Fiz. Tverd. Tela **38**, 2113 (1996) [Phys. Solid State **38**, 1165 (1996)].
[3] Z. S. Kachlishvili, Zh. Éksp. Teor. Fiz. **78**, 1955 (1980) [Sov. Phys. JETP **51**, 982 (1980)].
[4] Z. S. Kachlishvili and F. G. Chumburidze, Zh. Éksp. Teor. Fiz. **87**, 1834 (1984) [Sov. Phys. JETP **60**, 1055 (1984)].
[5] V. L. Bonch-Bruevich and S. G. Kalishnikov, *Physics of Semiconductors* (Nauka, Moscow, 1990).

Translated by W. J. Manthey

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Conductivity and persistent photoconductivity in GaAs epitaxial films containing single and double delta-doped layers

V. V. Valyaev, V. L. Gurtovoĭ, D. Yu. Ivanov, S. V. Morozov, V. V. Sirotkin,
Yu. V. Dubrovskiĭ, S. Yu. Shapoval, Yu. N. Khanin, E. E. Vdovin, and A. N. Pustovit

*Institute for Problems of Microelectronics and High-Purity Materials, Russian Academy of Sciences,
142432 Chernogolovka, Moscow Region, Russia*

(Submitted 21 January 1997)
Zh. Éksp. Teor. Fiz. **113**, 693–702 (February 1998)

Electrophysical parameters of single and double delta-doped layers in GaAs epitaxial films grown by the metal-organic chemical vapor deposition have been systematically investigated in the temperature range of 4.2 to 300 K. The 2D electron gas density distribution is affected by the overlap of wave functions in neighboring quantum wells, as a result of which the peak on the curve of the Hall mobility in the 2D electron gas versus the separation between the quantum wells shifts. The persistent photoconductivity in delta-doped layers is due to the change in the surface potential caused by the neutralization of the negative charge of surface states by photoexcited holes. A method for comparing delta-doped layers grown under different conditions at different depths from the sample surface has been suggested. © *1998 American Institute of Physics.* [S1063-7761(98)02402-0]

1. INTRODUCTION

Semiconducting structures with highly nonuniform doping of homogeneous semiconductors, when the dopant is concentrated in a very narrow layer (in the limiting case, in an atomic monolayer), are usually referred to as structures with delta-doped layers or simply delta-layers. The charge of doping impurities in the delta-layer generates a V-shaped potential well. As a result, the structure contains a two-dimensional (2D) electron gas. One can find in scientific journals several detailed reviews of 2D electron gas properties in delta-layers.[1–3]

In 1993 Zheng *et al.*[4] published their first report about the Hall mobility increase in GaAs structures with double delta-layers as compared to similar structures with single delta-layers in the temperature range of 77 to 300 K. Later a similar effect was observed in silicon-based structures.[5] On the other hand, in earlier publications an increase in the mobility in GaAs structures with periodic delta-doping at liquid-helium temperatures was reported.[6] This increase in the mobility is attributed to the redistribution of the electron density in real space in higher size-quantized subbands owing to the overlap of wave functions in neighboring quantum wells. This redistribution reduces the electron scattering due to ionized impurities in the doping planes in the middle of the quantum wells. The large spread in the results obtained by different researchers who measured mobilities in both double and single delta-layers was probably caused by different real widths of doping layers obtained in the process of growth.[7]

On the other hand, it is well known[8,9] that at a fixed doping level both the concentration of 2D electrons in a delta-layer in GaAs and their mobility strongly depend on the separation between the layer and surface. This is due to partial depletion of delta-layers by the negative surface potential. As a result, the potential well becomes asymmetrical, which leads to a shift of the ‘‘centers of mass’’ of electron wave functions and a redistribution of electrons among different size-quantized subbands. Therefore delta-layers with equal dopant concentrations but grown at different depths are not equivalent, and it is incorrect to compare their properties considering them as functions of growth conditions only.

The next important experimental fact is that persistent photoconductivity occurs in structures with delta-layers at temperatures below 100 K.[10] Several models of the persistent photoconductivity have been discussed in literature. It is ascribed to photoionization of DX-centers in GaAs resonant with the conduction band in the delta-layer,[11] which can lead to splitting of photoionized electron–hole pairs by the electric field generated by the delta-layer so that electrons remain in the potential well and holes go into the bulk of the material.[12] This model, which ignores both the surface and semi-insulating substrate, predicts a logarithmically long decay time of minority carriers due to the separation of nonequilibrium electrons and holes. A mechanism of persistent photoconductivity assuming photogeneration of a conducting hole channel parallel to the delta-layer in the semiconductor surface region was also suggested.[10]

All these facts indicate that no consistent approach to characterization of delta-layers has yet been developed and certain problems arise when measurements of electronic parameters of delta-layers obtained by different research groups are compared.

In the present work we have conducted systematic measurements of Hall concentrations and mobilities for single and double delta-layers in GaAs at temperatures ranging from 4.2 K to room temperature. As in earlier experiments,[4] we have detected in structures with double layers a peak on the curve of mobility as a function of separation between layers at liquid-nitrogen and room temperatures. At the liquid-helium temperature, however, the mobility is found to

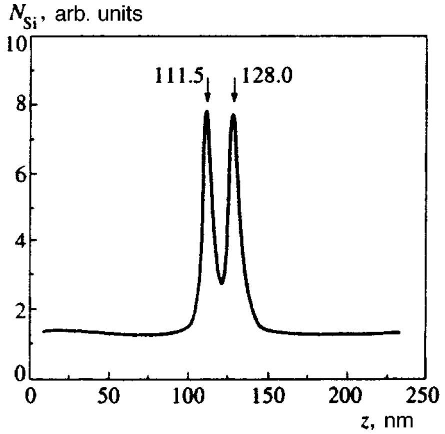


FIG. 1. Profile of silicon donor concentration measured using secondary-ion mass-spectroscopy; z is the depth with respect to the sample surface.

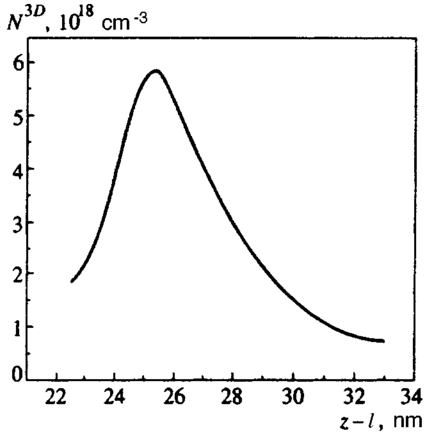


FIG. 2. Profile of bulk donor concentration N^{3D} derived from $C-V$ characteristics measured on satellite samples; $z-l$ is the distance from the sample surface. Before deposition of the Schottky barrier a surface layer of thickness $l \approx 75$ nm was etched off.

be independent of the separation between layers within the experimental uncertainty. This effect is due to the strong dependence of the potential well shape and conditions for the overlapping of wave functions in neighboring wells on temperature. We have also demonstrated directly in an experiment with a biased gate that persistent photoconductivity is at least partially controlled by the notable drop in the surface potential after exposure to light due to annihilation of photoexcited holes and electrons that were previously in equilibrium and occupied surface states. On the base of the analysis of experimental data, we have come to a conclusion that comparison of properties of different delta-layers is sensible at low temperature in the persistent conductivity regime, when the effect of surface potentials on the quantum well shape is reduced considerably.

2. SAMPLES

GaAs epitaxial films containing silicon delta-doped layers were grown by metal-organic chemical vapor deposition at a temperature of 600 °C on semi-insulating GaAs substrates. The total width of the lightly doped n^--GaAs epitaxial layer $[n=(1-2)\times 10^{15}$ cm^{-3} was 600 nm plus the separation d between the layers. The separation between the delta-layers and surface was $z=100$ nm. A doping level in delta-layers was $N_D=3\times 10^{12}$ cm^{-2}. In structures with double delta-layers the distance between the two layers was 8 to 48 nm.

The fabricated delta-layers were tested using the secondary-ion mass spectroscopy and measurements of capacitance–voltage characteristics. A typical silicon concentration profile in a structure containing delta-layers derived from secondary-ion mass-spectroscopic measurements is shown in Fig. 1. The measured FWHM of the donor distribution was 10 nm, whereas the spatial resolution provided by the technique was about 7 nm. The latter was estimated by measuring the profile of Al concentration around the heterojunction in a GaAs/AlGaAs structure performed in the same regime of the secondary-ion mass spectroscopy as in measurements of silicon concentration in delta-layers. This method of estimating the spatial resolution of silicon concentration measurements is not precise; nonetheless, it yields a realistic estimate of the FWHM in a delta-layer. The estimated half-width of the silicon donor distribution in delta-layers is 7 nm.

A typical donor concentration profile derived from measurements of capacitance–voltage characteristics $(C-V)$ is given in Fig. 2. Measurements were performed on satellite samples fabricated concurrently with the main samples but on conducting n^+-GaAs substrates. For measuring donor distributions by this method, we deposited aluminum or gold Schottky barriers on epitaxial film surfaces. The FWHM of impurity concentration profiles derived from $C-V$ characteristics was 4 nm. It is known that the width of the concentration profile derived from $C-V$ measurements is determined by the spatial width of the ground state wave function in a V-shaped potential well.[13] The width of the obtained $C-V$ profile depends weakly on the dopant distribution width below 5 nm. Published data on calculations of $C-V$ concentration profiles of delta-layers[13] as functions of the dopant distribution width have allowed us to set the latter parameter at 5–6 nm, which is in fair agreement with secondary-ion mass-spectroscopic measurements and, as will be shown below, with measurements of the Hall mobility in structures with double delta-layers.

3. MEASUREMENTS OF HALL PARAMETERS AND DISCUSSION

The Hall mobility and concentration were measured in both Hall bridge samples and Van-der-Pauw geometry. At room and liquid-nitrogen temperatures, measurements were performed in magnetic fields of up to 0.4 T, and at the liquid-helium temperature of 4.2 K in fields of up to 8 T. We also measured Hall parameters of some samples over the temperature range from 4.2 to 300 K, but these measurements did not yield any additional information of fundamental importance, so they are not discussed in this paper. Note only that the measurements of the Hall mobility in the tested samples in the temperature range of 4.2 to 300 K demonstrated a peak at temperatures around that of liquid nitrogen, just as in homogeneously doped bulk semiconductors. All

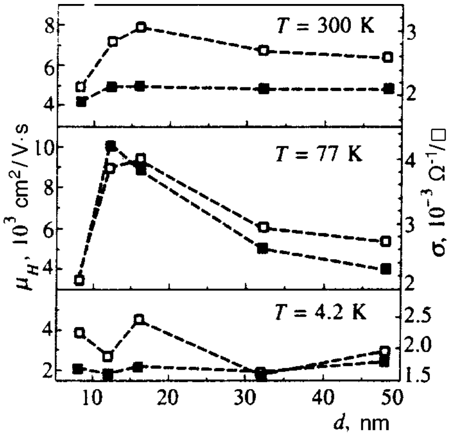


FIG. 3. Hall mobility μ_H (open squares) and conductivity σ (filled squares $\Omega^{-1}/\square$) as functions of separation d between delta-layers.

measurements were performed in the dc mode at a current across samples of 100 μA. Testing measurements demonstrated that this current did not heat up the electronic system even at liquid-helium temperatures.

In all samples, the Hall mobility at 300 K was in the range 2000–3000 cm^2/V·s, which is close to the mobility of the best GaAs samples with delta-layers grown by molecular-beam epitaxy.[14]

Figure 3 shows measurements of the conductivity and Hall mobility of epitaxial structures containing double delta-layers as functions of the separation between layers at different temperatures. Note the main features of these curves. The Hall mobility versus separation between layers has a broad peak at room temperature, a more pronounced peak at liquid-nitrogen temperature, and is rather flat at liquid-helium temperature. These data refer to measurements in the dark, i.e., no persistent photoconductivity was generated.

At this point, it is appropriate to make several remarks on measurements of the Hall concentration and mobility in structures with delta-layers. In structures with a doping level of 3×10^{12} cm^{-2} (the measured Hall concentration is set at 2.5×10^{12} cm^{-2}) the number of occupied subbands at liquid-helium temperature is three.[15] As the temperature increases, higher subband are naturally partially populated owing to thermal activation.

In delta-layers with several populated subbands, the contribution of scattering on the ionized impurities in the middle of the potential well to the resistivity is notably higher than that due to the intersubband scattering.[9] The mobilities in structures with delta-doping are relatively small $(1000-5000\,\mathrm{cm^2/V\cdot s})$, so transport properties of the electron gas in the delta-layer are usually described by assuming that each subband contains electrons with concentration n_i and mobility μ_i, and the conductivity σ is determined by adding conductivities due to electrons from different subbands:

$$\sigma=e\sum_i n_i\mu_i . \tag{1}$$

Note that, if such a description is applied to 2D structures with high mobilities, where the intersubband scattering is important, it is a very rough approximation.[16]

In experiments the measured parameters are the Hall concentration n_H and Hall mobility μ_H, which are related to the conductivities and populations in the subbands by the following expressions:

$$n_H=\frac{(\Sigma n_i\mu_i)^2}{\Sigma n_i\mu_i^2}, \tag{2}$$

$$\mu_H=\frac{\Sigma n_i\mu_i^2}{\Sigma n_i\mu_i}. \tag{3}$$

It follows that measurements of the Hall mobility yield rather inaccurate information concerning properties of the electronic system. The more appropriate parameter is conductivity. In our experiments the curves of conductivity and Hall mobility versus separation between layers are similar, so the observed effects can be discussed in terms of increasing mobility.

As in Ref. 4, we observed in our experiments a peak on the curve of mobility versus separation between the delta-layers at liquid-nitrogen and room temperatures. The mobility peak at room temperature was centered at $d_0\approx13$ nm. This fact is attributed to the overlap of wave functions in neighboring quantum wells, primarily those of the higher subbands. In this case the probability of finding electrons in the region between the layers is considerably higher. Electrons are farther from scattering impurities in comparison with the case of a single delta-layer, which leads to an increase in both mobility and conductivity.[4]

When the separation between layers is larger than d_0, the overlap between the wave functions is smaller, and at separations smaller than d_0 it becomes comparable to the width of the dopant distribution which causes the peak on the mobility curve.

It turned out, however, that at 4.2 K the conductivity and mobility are independent of the separation d between the layers. This observation can be qualitatively interpreted as follows. The depth and width of the V-shaped well depend sensitively on the temperature.[17] The lower the temperature, the narrower the potential well, so the peak on the curve of mobility versus separation between layers should shift toward smaller separations as the temperature decreases. In the studied structures, the peak was detected at liquid-nitrogen temperature at 13 nm, whereas the width of the dopant distribution was, as noted above, 5 to 7 nm. Therefore the overlap of the wave functions takes place when the dopant distributions in the quantum wells also overlap considerably. Moreover, at liquid-nitrogen temperature the contribution from electrons thermally excited to higher two-dimensional subbands can be quite considerable, and their wave functions overlap at larger distances. The latter statement is corroborated by the presence of the long tail of mobility on the side of larger separations at the liquid-nitrogen temperature.

4. PERSISTENT PHOTOCONDUCTIVITY

In all structures studied, the persistent photoconductivity was observed at the liquid-nitrogen and liquid-helium temperatures. In structures with single delta-layers, the resistivity decreased by 10–20% after illumination at $T=77$ K and by 30–40% at the liquid-helium temperature. In double delta-layers the persistent photoconductivity was essentially smaller and equal to $\sim 5\%$ of the dark conductivity.

As was noted above, photoionization of DX-centers, whose energy is lower than the Fermi level, can be one of the causes of the persistent photoconductivity.[11] It is well known,[18] however, that DX-centers in GaAs do not resonate with conduction electrons when the donor concentration in a single delta-layer is $N_D = 3\times 10^{12}$ cm^{-2}, i.e., their energy is above the Fermi level since the DX-center energy is 150–300 meV above the conduction band bottom. Nonetheless, in the studied samples with the Hall concentration of electrons in the layer $n = 2.5\times 10^{12}$ cm^{-2} ($N_D = 3\times 10^{12}$ cm^{-2}) persistent conductivity is observed at liquid-nitrogen temperatures and lower. The model suggested by Sanchez–Dehesa *et al.*[12] for interpreting the persistent photoconductivity does not take into account the surface potential, which should play an important role in thin epitaxial GaAs films (in GaAs the surface potential determined by charged surface states ranges between 0.7 and 0.8 V, depending on the surface processing technique[19]). Emergence of a hole layer near the surface[10] seems very improbable because it requires a surface potential of an unknown nature which should remain constant both under illumination and when the light is off.

The following cause of the persistent photoconductivity is more realistic in our opinion. Owing to the presence of the surface potential in GaAs, the delta-layer that is closer to the surface is slightly depleted. There is a constant electric field between the delta-layer and the surface in the absence of additional doping:

$$E = \varphi_s / z,$$

where φ_s is the surface potential and z is the depth at which the delta-layer is located. The negative charge of the surface states is

$$Q_s = e\left(N_D - \sum_i n_i \right),$$

where N_D is the two-dimensional concentration of donors in the delta-layer. Under interband illumination, electron–hole pairs generated in the region of the surface electric field split up. Electrons are driven to the delta-layer, and holes annihilate with electrons in surface states. The well-known effect of band flattening under illumination takes place. When the light is off, electrons from the delta-layer partially occupy emptied surface states. The full recovery, however, is impossible because of the potential barrier created between the layer and surface. If the barrier cannot be tunneled through, it can be passed over only by thermally activated electrons, therefore, the lower the temperature, the larger the relative decrease in the delta-layer resistance in comparison with the dark resistance, as was observed in experiments.

In order to check out this assumption, we performed a simple experiment on gated samples with single delta-layers. The idea of the experiment was that the charge of surface states could be regenerated owing to the delta-layer electrons driven to surface states via an external electric circuit if the suggested model is correct. A diagram of the experimental configuration is shown in the insert to Fig. 4. Measurements were performed at 4.2 K. The sample was cooled in the dark. The delta-layer resistance was measured with a ‘‘hung’’ gate (the region on the left of t_0 in Fig. 4). It was 600 Ω/□. Then a filament lamp placed in the cryostat next to the sample was switched on for some time. Owing to the photoconductivity (during the time interval from t_0 to t_1, Fig. 4) the measured resistivity was 300 Ω/□. Then the illumination was turned off. Within a few seconds the resistivity increased to 400 Ω/□ and remained constant. The persistent photoresistivity remained constant to within 0.5% for 24 h (between t_1 and t_2). Then the gate was connected to one of the ohmic contacts, and as a result, the resistivity increased to 520 Ω/□, although it did not return to the initial value (the region on the right of t_2 in Fig. 4). After warming up and subsequent cooling in the dark, the initial resistivity of the sample was restored.

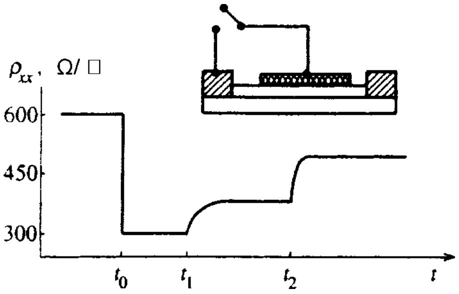


FIG. 4. Resistivity of a Hall-bridge structure with an aluminum gate versus time (the experimental configuration is shown in the insert) before ($t<t_0$), during ($t_0<t<t_1$) and after ($t_1<t<t_2$) illumination, and after filling surface states with photogenerated electrons ($t>t_2$). The time scale on different sections of the curve is different.

It is impossible to restore the initial sample resistivity by connecting the gate to the delta-layer because similar processes associated with optical recharging of deep levels in the interface region or in the interior of the semi-insulating GaAs should also take place on the side of the semi-insulating substrate. Note that the nontransparency of the gate had little effect because neutralization of the negative surface charge due to illumination near gate edges led to the total neutralization of charges under the metal gate because surface states are separated from the gate only by a thin tunnel layer of natural GaAs oxide.

In testing experiments, the sample was illuminated by a lamp through a GaAs filter in order to cut off the interband light. In this case no photoconductivity was detected.

Fu and Willander[20] showed that the electron mobility in the delta-layer increased with the surface potential generated, for example, by changing the gate voltage. This was attributed to a shift of the wave functions away from scattering impurities in the delta-layer (Fig. 5).

In our experiments, however, we detected an opposite effect. For example, in structures with a single delta-doped layer the Hall mobility in the dark was 2900 cm^2/V·s and in the presence of persistent photoconductivity it was 4100 cm^2/V·s.

This controversy can be tentatively explained as follows.

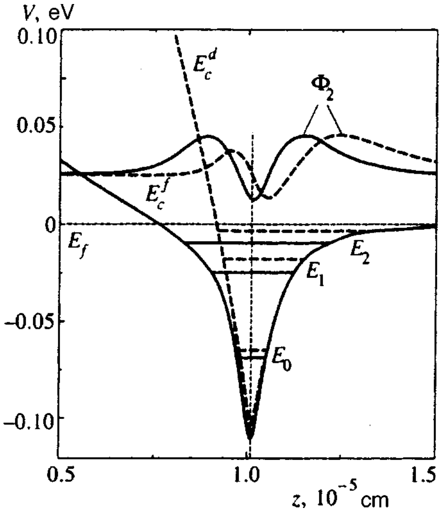


FIG. 5. Self-consistent calculation of the bands in the potential well around the delta-layer: E_c^f is the conduction band edge in the regime of persistent photoconductivity for $\varphi_s = 0.1$ V (solid curve); E_c^d is the conduction band edge in darkness with $\varphi_s = 0.8$ V (dashed curve); E_0, E_1, and E_2 are energies of size-quantized levels; solid lines show energies in the persistent conductivity regime, and dashed lines plot energies in darkness. All energies are measured with respect to the Fermi level, $E_F = 0$; z is the distance to the surface. Concentration of donors in the delta-layer is $N_D = 3\times10^{12}$ cm^{-2}, the temperature in calculations is $T = 4$ K. For the level E_2 the graph shows the wave functions Φ_2: the solid curve is the wave function in the persistent photoconductivity regime, the dashed line shows the wave function without illumination. The curves of the wave functions are shifted along the vertical axis, their vertical scale is arbitrary. It is clearly seen that in the dark regime the wave function is farther from the doping plane shown by the vertical dashed line.

Our calculations indicate that the electron concentration increases by about 20% when the surface potential is reduced. Eighty-five percent of them go to higher subbands, where the mobility is a factor of about five higher than that in the ground-state subband.[16] In our model calculations we used the values of the surface potential $\varphi_s = 0.1$ V in the presence of persistent photoconductivity and $\varphi_s = 0.8$ V for dark conductivity. This redistribution of electrons is attributed to the fact that the potential well becomes wider and more symmetrical when the surface potential is lowered (Fig. 5). Simple estimates by Eq. (3) yield the 20% increase in the Hall mobility, which is in a fair agreement with experimental data.

An important point is that two processes making opposite contributions to the resulting Hall mobility occur concurrently when the surface potential increases: (1) electrons are repelled from the plane of scatterers, which should lead to a higher electron mobility, primarily in the higher subbands, where the wave function shift is larger;[20] (2) elevation of higher subbands in a more narrow asymmetrical potential well, as a result of which the population of higher subbands with higher mobilities decreases. Our data indicate that the second effect dominates, at least in the studied structures, and the Hall mobility drops with the surface potential.

The investigation of the nature of persistent photoconductivity has shown, in particular, that it is incorrect to compare properties of double delta-layers with those of single layers because of the presence of the surface potential.

5. CONCLUSIONS

In double delta-layers the conductivity is higher because of the overlap of the wave functions and the displacement of electrons away from scatterers. This effect, however, depends on temperature and broadening of the dopant distribution in delta-layers. The persistent photoconductivity in epitaxial GaAs films containing delta-layers at low temperatures is due to flattening of bands outside the potential well after exposure to interband light. Comparison between properties of delta-layers grown under different conditions at different depths from the sample surface is possible only in the presence of persistent photoconductivity at low temperatures and small surface potential. The Hall conductivity in delta-layers depends on the surface potential (gate voltage) as a result of the competition among various processes caused by changes in the surface potential. The Hall mobility in delta-layers with widths of the donor distribution of 5 to 6 nm in the vertical direction is essentially smaller at high surface potentials and low temperatures.

The work was supported by the Russian Fund for Fundamental Research, Project No. 95-02-06328.

[1] K. Ploog, J. Cryst. Growth **81**, 304 (1987).
[2] E. F. Shubert, in *Semiconductor and Semimetals: Epitaxial Microstructures*, ed. by A. G. Gossard, Academic Press, New York (1994), Vol. 40, p. 1.
[3] A. Ya. Shik, Fiz. Tekh. Poluprovodn. **26**, 1161 (1992) [Sov. Phys. Semicond. **26**, 649 (1992)].
[4] X. Zheng, T. K. Carns, K. L. Wang *et al.*, Appl. Phys. Lett. **62**, 504 (1993).
[5] H. H. Radamson, M. R. Sardela, Jr., O. Nur *et al.*, Appl. Phys. Lett. **64**, 1842 (1994).
[6] T. Ihn, H. Kostial, R. Hey *et al.*, in *Extended Abstract of the 1992 International Conf. on Solid State Devices and Materials*, Tsukuba Center Building, Tsukuba, Japan (1992), p. 313.
[7] P. M. Koenraad, A. C. L. Heessels, F. A. P. Blom *et al.*, Physica B **184**, 221 (1993).
[8] T. Makimoto, N. Kobayashi, and Y. Horikoshi, J. Appl. Phys. **63**, 5023 (1988).
[9] G. Q. Hai and N. Studart, Phys. Rev. B **52**, 2245 (1995).
[10] A. G. de Olivera, G. M. Ribeiro, D. A. W. Soares *et al.*, J. Appl. Phys. **78**, 2659 (1995).
[11] S. Arscott, M. Missous, and L. Dobaczewski, Semicond. Sci. Technol. **7**, 620 (1992).
[12] J. Sanchez-Dehesa, D. Lavielle, E. Ranz *et al.*, Semicond. Sci. Technol. **6**, 445 (1991).
[13] E. F. Shubert, J. M. Kuo, and R. F. Kopf, J. Electron. Mater. **19**, 521 (1990).
[14] P. M. Koenraad, F. A. P. Blom, C. J. G. M. Langerak *et al.*, Semicond. Sci. Technol. **5**, 861 (1990).
[15] G. Q. Hai, N. Studart, and F. M. Peeters, Phys. Rev. B **52**, 8363 (1995).
[16] W. de Lange, F. A. Blom, P. J. van Hall, *et al.*, Physica B **184**, 216 (1993).
[17] I. A. Panaev, S. A. Studenikin, D. I. Lubyshev, and V. P. Migal', Semicond. Sci. Technol. **8**, 1822 (1993).
[18] R. C. Newman, Semicond. Sci. Technol. **9**, 1749 (1994); T. N. Theis, P. M. Mooney, and S. L. Wright, Phys. Rev. Lett. **60**, 361 (1988); D. C. Chadi and K. J. Chang, Phys. Rev. Lett. **61**, 873 (1988).
[19] E. H. Rhoderick, *Metal–Semiconductor Contacts*, Clarendon Press, Oxford (1978).
[20] Y. Fu, and M. Willander, J. Appl. Phys. **78**, 3504 (1995).

Translation provided by the Russian Editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Fine structure of localized exciton levels in quantum wells

S. V. Gupalov, E. L. Ivchenko,*) and A. V. Kavokin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

(Submitted 23 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 703–714 (February 1998)

A theory is constructed for the long-range exchange and retarding interactions between an electron and a hole in a quantum well. A method is developed that makes it possible to calculate the ground and excited states of an exciton localized as a whole on a width fluctuation of a quantum well in the form of a rectangular island. It is shown that taking into account the electron–hole interaction mechanisms considered here causes the radiation doublet of the exciton to split into two components polarized along the sides of the rectangle. The dependence of the magnitude and sign of this splitting on the linear dimensions of the island and the level number of the localized exciton are analyzed. © *1998 American Institute of Physics.* [S1063-7761(98)02502-5]

1. INTRODUCTION

In an ideal semiconductor heterostructure with a quantum well, the motion of the charge carriers (electrons or holes) is quantized in the growth direction of the structure and is free in the plane of the interfaces. In actual structures grown by even the most perfect molecular-beam epitaxy, the heteroboundary is not ideally flat and, in the best case, is a surface whose sections lie in two adjacent planes separated by a monomolecular layer. The topology of this surface depends on the materials of the heterojunction and the process conditions, in particular on the temperature and the time at which growth was terminated. If there is no correlation between the interfaces, the width of the well fluctuates, assuming one of three values. As a result, a quasi-two-dimensional carrier finds itself in a fluctuation potential whose amplitude is determined by the difference of the size-quantization energies in two ideal wells with widths that differ by one or two monolayers. In this potential, a tail of localized exciton states is formed, corresponding to the low-temperature photoluminescence of undoped structures with quantum wells.[1–3] In the papers that we know of, localized excitons were calculated for an axially symmetric Gaussian potential[4] or in a circular-island model.[5] In structures of the GaAs/AlGaAs(001) type, the state of such a localized exciton, optically active in the interface plane, is doubly degenerate. It is clear that axial symmetry of the localized potential is an exception and, as a rule, it is anisotropic in the interface plane. A decrease in the symmetry must remove the degeneracy of the radiation doublet and affect the polarization of the photoluminescence when the excitons are optically oriented.

A theory of excitons localized on anisotropic islands of a monolayer fluctuation of the width of a quantum well is constructed in this paper. It is assumed that the linear size of the island exceeds the Bohr radius of a quasi-two-dimensional exciton. It is shown that the splitting of the doublet caused by the long-range exchange interaction of the electron and hole in the exciton in this case is several tens of microelectron volts when the sides of the rectangular island differ by a factor of 1.5 or 2. Such splitting significantly exceeds the natural linewidth $\hbar/\tau$ of the radiation of a localized exciton, since a typical value for its lifetime is about 10^{-9} sec. Gammon *et al.*[6] recently studied the photoluminescence spectrum of a localized exciton from an individual island in a GaAs/AlGaAs(001) quantum well in the optical near-field regime and detected splitting of the $e1-hh1(1s)$ doublet into two components polarized along the [110] and [110] axes. In the photoluminescence excitation spectrum, the same authors observed excitation of a state of an exciton localized on an island as a single whole. For the ground and four excited states, the splitting $\Delta=E_{1\bar{1}0}-E_{110}$ between the sublevels polarized along the [110] and [110] directions equalled -25, $+41$, $+45$, -22, and -47 μeV, respectively. It was therefore of interest to elucidate whether it is possible to choose the size of the rectangular island so as to reproduce the indicated sequence of signs of the splitting Δ. Preliminary results of this work were presented in Ref. 7.

2. EXCHANGE AND RETARDING MECHANISMS OF ELECTRON–HOLE INTERACTIONS

A consistent theory of electron–hole exchange interaction in semiconductors has been developed by Pikus and Bir.[8,9] In the effective-mass approximation, the Coulomb-interaction operator between an electron and a hole in a semiconductor crystal includes three contributions, describing, respectively, the direct, or intraband, Coulomb interaction (U_C) and the long-range ($U_{\rm exch}^{\rm long}$) and short-range ($U_{\rm exch}^{\rm short}$) exchange interaction. We introduce the two-particle excited states of the crystal $|m,\mathbf{k}_e;n,\mathbf{k}_h\rangle$, where $\mathbf{k}_{e,h}$ is the wave vector of the electron or hole, and the indices m and n label the degenerate states of the electron in the conduction band and the hole in the valence band at $k_{e,h}=0$ (for definiteness, a direct semiconductor with cubic symmetry and an extremum at the Γ point is considered). Then the matrix elements of the U_C and $U_{\rm exch}^{\rm long}$ operators between these states can be written as

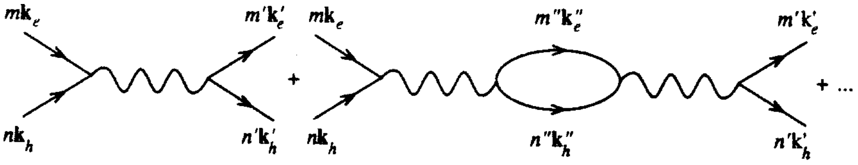


FIG. 1.

$$\langle m',\mathbf{k}'_e;n',\mathbf{k}'_h|U_C|m,\mathbf{k}_e;n,\mathbf{k}_h\rangle = -\frac{1}{V}\frac{4\pi e^2}{\kappa_0|\mathbf{k}_e-\mathbf{k}'_e|^2}\,\delta_{m'm}\delta_{n'n}\delta_{\mathbf{k}_e+\mathbf{k}_h,\mathbf{k}'_e+\mathbf{k}'_h}, \tag{1}$$

$$\langle m',\mathbf{k}'_e;n',\mathbf{k}'_h|U^{(\mathrm{long})}_{\mathrm{exch}}|m,\mathbf{k}_e;n,\mathbf{k}_h\rangle = \frac{1}{V}\frac{4\pi e^2\hbar^2}{\kappa_b m_0^2E_g^2}\frac{(\mathbf{K}\cdot\mathbf{p}_{m'\overline{n'}})(\mathbf{K}\cdot\mathbf{p}_{m\bar{n}})^*}{K^2}\,\delta_{\mathbf{k}_e+\mathbf{k}_h,\mathbf{k}'_e+\mathbf{k}'_h}. \tag{2}$$

Here $\mathbf{K}$ is the total wave vector, $\mathbf{k}_e+\mathbf{k}_h=\mathbf{k}'_e+\mathbf{k}'_h$; m_0 is the free-electron mass; E_g is the band gap; $\mathbf{p}_{m\bar{n}}$ is the matrix element of the momentum operator, calculated between the electronic Bloch functions $|m,\mathbf{k}=0\rangle$ and $|\bar{n},\mathbf{k}=0\rangle$ (the hole state $n,\mathbf{k}$ and the electron state $\bar{n},-\mathbf{k}$ are coupled to each other by the time-inversion operator); κ_0 and κ_b are the permittivities, low-frequency and high-frequency (at the electron–hole-pair excitation frequency); e is the charge of the electron; and V is the volume of the crystal. The interaction given by Eq. (2) is described by the sequence of diagrams in Fig. 1, where a pair of lines represents the Green's function of electron–hole excitation, while a wavy line represents the Fourier component of the unscreened Coulomb potential $4\pi e^2/V|\mathbf{k}_e-\mathbf{k}_h|^2$. This interaction can thus be interpreted as the result of the virtual recombination and generation of an electron–hole pair, and it is reduced by a factor of κ_b when the chain of virtual electron–hole excitations induced by the Coulomb potential is taken into account in the diagrams. Equation (2) can be obtained by one more method: the macroscopic electric field generated by the electron–hole pair can be calculated, and the self-consistent effect of this field on the energy of the pair can be taken into account.

The Fourier components of the Coulomb potential with wave vectors $\mathbf{b}+\mathbf{k}_{e'}-\mathbf{k}_e$, where $\mathbf{b}$ are nonzero reciprocal lattice vectors, contribute to the short-range interaction. When k_e and k_h are small enough to satisfy the criterion for the applicability of the effective-mass method, the operator $U^{\mathrm{short}}_{\mathrm{exch}}$ has the character of a contact interaction and can be represented in the form $\Delta_{m'n',mn}a_0^3\delta(\mathbf{r}_e-\mathbf{r}_h)$, where a_0 is the lattice constant and the factor a_0^3 is separated out so that the coefficients $\Delta_{m'n',mn}$ will have the dimension of energy. The dependence of these coefficients on the band indices is found from symmetry considerations, while their absolute magnitudes are determined by comparing theory with experiment by studying the fine structure of the exciton levels (see, for example, Refs. 10–12). The number of linearly independent coefficients coincides with the number of irreducible representations contained in the direct product $\Gamma_c\times\Gamma_v$ according to which the electron states at the bottom of the conduction band and at the top of the valence band transform. For illustration, let us consider the pair of bands Γ_6 and Γ_7 in GaAs-type semiconductors: $\Gamma_6\times\Gamma_7=\Gamma_2+\Gamma_5$. It is convenient to go to a basis of electron–hole excitations in which the three states $|\nu,\mathbf{k}_e,\mathbf{k}_h\rangle$ $(\nu=x,y,z)$ are optically active in the polarization $\mathbf{e}\parallel\nu$, while the optical transition to the fourth state $|\Gamma_2,\mathbf{k}_e,\mathbf{k}_h\rangle$ is forbidden. The exchange interaction given by Eq. (2) involves only the states $|\nu,\mathbf{k}_e,\mathbf{k}_h\rangle$ and has in the new basis the form

$$\langle\nu',\mathbf{k}'_e,\mathbf{k}'_h|U^{(\mathrm{long})}_{\mathrm{exch}}|\nu,\mathbf{k}_e,\mathbf{k}_h\rangle=\frac{4\pi}{\kappa_bV}\left(\frac{e\hbar p_0}{m_0E_g}\right)^2\frac{K_{\nu'}K_\nu}{K^2}\,\delta_{\mathbf{K}',\mathbf{K}}, \tag{3}$$

where p_0 is the interband matrix element of the momentum operator for the optical transition to the state $|\nu\rangle$.

We have generalized Pikus and Bir's theory,[8] taking into account the retarding interaction that results from the successive emission and absorption of a transverse photon by an electron–hole pair, described by the diagram in Fig. 2. Here the dashed line represents the photon Green's function in a medium with permittivity κ_b, while the vertex corresponds to the factor $-i/\hbar$ as well as the electron–photon interaction matrix element

$$-\frac{e}{m_0}\left(\frac{2\pi\hbar}{V\omega_K\kappa_b}\right)^{1/2}\mathbf{e}\cdot\mathbf{p}_{m\bar{n}}$$

for photon absorption or the complex-conjugate expression for photon emission, where the photon frequency is $\omega_K=(c/\sqrt{\kappa_b})K$, and $\mathbf{e}$ is its polarization vector. As a result, we get

$$\Sigma_{\mathrm{el-phot}}(\nu',\mathbf{k}'_e,\mathbf{k}'_h;\nu,\mathbf{k}_e,\mathbf{k}_h;\omega)=-\frac{4\pi}{\kappa_bV}\left(\frac{e\hbar p_0}{m_0E_g}\right)^2 \times\left(\delta_{\nu\nu'}-\frac{K_\nu K_{\nu'}}{K^2}\right)\frac{k^2}{K^2-k^2-i0}\,\delta_{\mathbf{K}',\mathbf{K}}, \tag{4}$$

for the matrix elements of the retarding interaction, where $k=\sqrt{\kappa_b}(\omega/c)$ and ω is the excitation frequency of the electron–hole pair or, in rigorous diagram language, the argument of the self-energy function that appears when one calculates the Green's function of an electron–hole pair, and the product $\omega\omega_K$ is replaced with $(E_g/\hbar)^2$ in the common factor. The real and imaginary parts of $\Sigma_{\mathrm{el-phot}}$ determine, respectively, the renormalization of the energy and the damping of the pairwise excitation caused by photon emis-

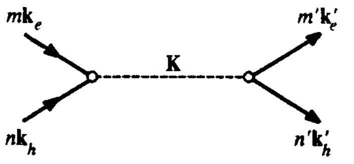


FIG. 2.

sion. In deriving Eq. (4) we took into account that the matrix $e_{1,\nu'}e_{1,\nu}+e_{2,\nu'}e_{2,\nu}$ (where $\mathbf{e}_1\perp\mathbf{e}_2\perp\mathbf{K}$) is the projection operator onto a plane perpendicular to the vector **K**.

An exciton calculated taking into account only the main interaction given by Eq. (1) is called a mechanical exciton.[13] A Coulomb exciton is found neglecting the retardation: according to Eq. (3), the splitting between the longitudinal and transverse states of a $1s$ exciton is associated with the microscopic parameter p_0 by the relationship

$$\hbar\omega_{LT}=\frac{4}{\kappa_b a_B^3}\left(\frac{e\hbar p_0}{m_0 E_g}\right)^2, \tag{5}$$

where a_B is the exciton Bohr radius. Finally, taking the retarding interaction into account results in transverse exciton polaritons. When we have $K\sim k$, the splitting of the polariton branches substantially exceeds the longitudinal–transverse splitting $\hbar\omega_{LT}$, and the polariton effect can be neglected by comparison with the long-range exchange interaction only when $K\gg k$.

3. ELECTRON–HOLE INTERACTION IN A QUANTUM WELL

Let us consider electron–hole excitations in a semiconductor quantum well in which the single-frequency states are characterized by the size-quantization subband number, by the two-dimensional wave vector $\mathbf{k}^{\parallel}=(k_x,k_y)$, and by spin indices m,n, running over two values. In quantum wells grown using GaAs-type semiconductors, we have $m=\pm 1/2$ for an electron and $n=\pm 3/2$ for a heavy hole or $n=\pm 1/2$ for a light hole. If intersubband mixing is neglected, the envelope of the two-frequency wave function has the form

$$\Psi(\mathbf{r}_e,\mathbf{r}_h)=S^{-1}\exp[i(\mathbf{k}_e^{\parallel}\cdot\boldsymbol{\rho}_e+\mathbf{k}_h^{\parallel}\cdot\boldsymbol{\rho}_h)]\varphi_e(z_e)\varphi_h(z_h). \tag{6}$$

Here $\boldsymbol{\rho}$ is the component of three-dimensional vector $\mathbf{r}$ in the interface plane (x,y), S is the area of the sample, and $\varphi_e(z_e)$ and $\varphi_h(z_h)$ are single-frequency size-quantization functions. In a symmetrical quantum well, these functions are characterized by a definite parity with respect to reflection in a plane passing through the center of the well, $z=0$. For definiteness, we shall assume in what follows that $\varphi_{e,h}(-z)=\varphi_{e,h}(z)$.

Let us calculate the matrix elements of the long-range exchange and retarding interactions between a quasi-two-dimensional electron and hole. To do this, it is necessary to multiply the sum of the matrix elements of the operators $U_{\text{exch}}^{\text{long}}$ and $U_{\text{el-phot}}$ by

$$\exp[i(K_z z-K_z' z')]\varphi_e(z')\varphi_h(z')\varphi_e(z)\varphi_h(z),$$

to sum over $K_{z'}$ and K_z, and to integrate over z' and z. Taking into account the symmetry of the functions $\varphi_{e,h}(z)$, we get

$$\Sigma(m',n',\mathbf{k}_e',\mathbf{k}_h';m,n,\mathbf{k}_e,\mathbf{k}_h;\omega)$$

$$=\frac{4\pi}{S\kappa_b}\left(\frac{e\hbar}{m_0E_g}\right)^2\delta_{\mathbf{k}_e'+\mathbf{k}_h',\mathbf{k}_e+\mathbf{k}_h}\Bigg\{[(\mathbf{q}\cdot\mathbf{p}_{m'\overline{n'}})(\mathbf{q}\cdot\mathbf{p}_{\overline{n}m})$$

$$-(\mathbf{p}_{m'\overline{n'}}\cdot\mathbf{p}_{\overline{n}m})k^2]\frac{iP(\gamma)}{2\gamma}+p^z_{m'\overline{n'}}p^z_{\overline{n}m}$$

$$\times\left[Q+\frac{i}{2}\gamma P(\gamma)\right]\Bigg\}, \tag{7}$$

where we have written $\gamma=\sqrt{k^2-q^2}$ and $k^2=(\omega/c)^2\kappa_b$ as before, the superscript $\parallel$ on the two-dimensional vectors $\mathbf{k}_e$ and $\mathbf{k}_h$ is omitted, $\mathbf{q}$ is a two-dimensional vector with components $(k_{e,x}+k_{h,x},k_{e,y}+k_{h,y})$, and

$$P(\gamma)=\int dz'\int dz\,\varphi_e(z')\varphi_h(z')\varphi_e(z)\varphi_h(z)$$

$$\times\exp(i\gamma|z-z'|),\quad Q=\int dz\,\varphi_e^2(z)\varphi_h^2(z).$$

The separate contribution of the long-range exchange interaction is obtained from Eq. (7) by formally going to the limit $c\to\infty$, i.e., $k\to 0$ or $\gamma=\sqrt{k^2-q^2}\to iq$. The diagonal components of Σ become real in this case, since their imaginary parts are associated exclusively with the photon-emission process. An expression for the matrix elements of the long-range exchange interaction between an electron and a hole in a quantum well was obtained earlier in Ref. 14 and is represented in a more unwieldy form, containing not the envelope $\varphi_{e,h}(z)$ but their Fourier components

$$C(k_z)=\int dz\,\exp(-ik_z z)\varphi(z).$$

Neglecting intersubband mixing, we write the envelope of the wave function of a free exciton in an ideal quantum well in the form

$$\Psi_{\text{exc}}(\mathbf{r}_e,\mathbf{r}_h)=S^{-1/2}\exp(i\mathbf{q}\cdot\mathbf{R}_\parallel)f(\boldsymbol{\rho})\varphi_e(z_e)\varphi_h(z_h), \tag{8}$$

where $\boldsymbol{\rho}=\boldsymbol{\rho}_e-\boldsymbol{\rho}_h$, and $\mathbf{R}_\parallel$ is the center of gravity of the exciton in the xy plane. The fine structure and the dispersion of free quasi-two-dimensional excitons in quantum wells, taking into account exchange interaction and retardation (or the polariton effect), has been calculated in a number of papers.[14–17] Using Eq. (7), we get, in accordance with Refs. 14–17,

$$E_1=-\frac{2\pi}{\kappa_b}i\left(\frac{e\hbar p_0}{m_0E_g}\right)^2\frac{k^2}{\sqrt{k^2-q^2}}P(\sqrt{k^2-q^2})f^2(0).$$

$$E_2=-\frac{2\pi}{\kappa_b}i\left(\frac{e\hbar p_0}{m_0E_g}\right)^2\sqrt{k^2-q^2}P(\sqrt{k^2-q^2})f^2(0),$$

$$E_3=\frac{2\pi}{\kappa_b}i\left(\frac{e\hbar p_0'}{m_0E_g}\right)^2\left[Q-\frac{i}{2}\sqrt{k^2-q^2}P(\sqrt{k^2-q^2})\right]f^2(0), \tag{9}$$

for the dispersion of the $1s$ exciton with a heavy $(hh1)$ or a light $(lh1)$ hole, where $E_\alpha=E_\alpha'-iE_\alpha''$ is the complex energy, the real part of which equals $E_\alpha'=\hbar(\omega-\omega_0)$ ($\hbar\omega_0$ is the energy of a mechanical exciton at $\mathbf{q}=0$), while the imaginary part determines the radiative lifetime of the exciton, $\tau=2\hbar/E_\alpha''$; the subscript $\alpha=1,2,3$ indicates the polarization of the exciton; axis 1 is perpendicular to the plane that contains the z axis and wave vector $\mathbf{q}$, axis 2 is parallel to $\mathbf{q}$, axis 3 is parallel to the z axis, $p_0(e1-hh1)=p_{cv}$, $p_0(e1-lh1)=p_{cv}/\sqrt{3}$, $p_0'(e1-hh1)=0$, $p_0'(e1-lh1)$

$=2p_{cv}/\sqrt{3}$, and p_{cv} is the interband matrix element $i\langle S|p_z|Z\rangle$ in direct semiconductors with the zincblende structure.

For excitons with a heavy hole, we go to two basis states $|\nu,\mathbf{k}_e,\mathbf{k}_h\rangle$, polarized along the fixed axes x,y, and two optically inactive states with total spin $m+n=\pm 2$. In this basis, Eq. (7) takes the form

$$\begin{aligned}\Sigma(\nu',\mathbf{k}_e',\mathbf{k}_h';\nu,\mathbf{k}_e,\mathbf{k}_h;\omega)&=\frac{1}{Sf^2(0)}\left[E_1\left(\delta_{\nu'\nu}-\frac{q_{\nu'}q_\nu}{q^2}\right)\right.\\&\quad\left.+E_2\,\frac{q_{\nu'}q_\nu^*}{q^2}\right]\delta_{\mathbf{k}_e'+\mathbf{k}_h',\mathbf{k}_e+\mathbf{k}_h}\\&=\frac{2\pi}{S\kappa_b}\,i\left(\frac{e\hbar p_0}{m_0E_g}\right)^2\frac{q_{\nu'}q_\nu-k^2\delta_{\nu'\nu}}{\sqrt{k^2-q^2}}\\&\quad\times P(\sqrt{k^2-q^2})\,\delta_{\mathbf{k}_e'+\mathbf{k}_h',\mathbf{k}_e+\mathbf{k}_h}.\end{aligned}\tag{10}$$

4. LOCALIZED EXCITON STATES IN THE MODEL OF INFINITELY HIGH BARRIERS

Let us consider the energy and exchange splitting of an exciton localized at the width fluctuations of a quantum well. Beyond the limits of the island, the quantum well consists of N monomolecular layers (in GaAs, the width of a monolayer is $a_0/2=2.8$ Å). One of the interfaces of the barrier is flat, while the second interface in the region of the island shifts by a monolayer in the depth of the barrier; i.e., the width of the well equals $(N+1)a_0/2$ in this region. The island is chosen in the shape of a rectangle oriented along the x and y axes.

For an estimate, we shall first carry out the treatment in terms of a simple model of infinitely high barriers and neglect the spread of the exciton beyond the limits of the island. Assuming that the linear dimensions L_x and L_y of the rectangle exceed the exciton Bohr radius a_B, we write the envelope of the exciton wave function in the form

$$\Psi_{\mathrm{exc}}(\mathbf{r}_e,\mathbf{r}_h)=F(X,Y)f(\boldsymbol{\rho})\varphi_e(z_e)\varphi_h(z_h),\tag{11}$$

where f, φ_e, and φ_h are the functions introduced into Eq. (8) and characterize the state of a free exciton in an ideal well of width $(N+1)a_0/2$, X and Y are the components of vector $\mathbf{R}_\parallel$, and the function $F(X,Y)$ describes the localization of the 1s exciton as a whole in the plane of the interfaces. Using as boundary conditions the vanishing of this function at the perimeter of the rectangle, we get a set of solutions in the form

$$F(X,Y)=F_x(X)F_y(Y),\tag{12}$$

where $F_\alpha(x_\alpha)$ $(\alpha=x,y)$ are one-dimensional size-quantization functions

$$\sqrt{\frac{2}{L_\alpha}}\begin{cases}\cos(\pi j x_\alpha/L_\alpha) & \text{for odd } j,\\ \sin(\pi j x_\alpha/L_\alpha) & \text{for even } j.\end{cases}\tag{13}$$

The energy of the localized exciton, measured from the bottom of the exciton band in a $(N+1)$-monolayer well, is characterized by the two indices j and j':

$$E_{jj'}=\frac{\hbar^2}{2M}\left[\left(\frac{j}{L_x}\right)^2+\left(\frac{j'}{L_y}\right)^2\right],\tag{14}$$

where M is the translational mass of the exciton as it moves in the xy plane. Each level $E_{jj'}$ consists of four sublevels, which, for a heavy exciton in GaAs/AlGaAs-type quantum wells, are characterized by the projection of the total spin $m+n=\pm 1,\pm 2$. Let us choose the basis states in the form $|j,j',x\rangle$, $|j,j',y\rangle$, and $|j,j',\pm 2\rangle$, where $|j,j',x\rangle=(|j,j',1\rangle+|j,j',-1\rangle)/\sqrt{2}$, and $|j,j',y\rangle=(|j,j',1\rangle-|j,j',-1\rangle)/\sqrt{2}$. For normal incidence of the light, the matrix element of the optical excitation of an exciton is proportional to the integral $\int dXdYF(X,Y)$. Sublevels $E_{jj',x}$ and $E_{jj',y}$ with odd j,j' are therefore optically active in the polarizations $\mathbf{e}\parallel x$ and $\mathbf{e}\parallel y$, respectively, while all other exciton states $|j,j',\pm 2\rangle$ with any j,j' and $|j,j',x\rangle$, $|j,j',y\rangle$ with at least one even index j or j' are optically inactive. However, the long-range exchange interaction given by Eq. (10) causes splitting between sublevels $E_{jj',x}$ and $E_{jj',y}$ for arbitrary j and j'. Specifically, going from the functions $F(x)$ to the Fourier components

$$F(q)=\int dx\,F(x)e^{-iqx}$$

and using Eq. (10) for the self-energy function Σ, we find the renormalization of the energy of localized exciton $|j,j',\nu\rangle$ $(\nu=x,y)$:

$$\begin{aligned}\Delta E_\nu=&\frac{2\pi}{S\kappa_b}\left(\frac{e\hbar p_0}{m_0E_g}\right)^2\sum_{\mathbf{q}}\frac{q_\nu^2-k^2}{\sqrt{q^2-k^2}}F_x^2(q_x)F_y^2(q_y)\\&\times[\theta(q-k)P(\sqrt{k^2-q^2})-i\theta(k-q)\mathrm{Im}\,P(\sqrt{k^2-q^2})],\end{aligned}\tag{15}$$

where $\theta(t)=1$ holds for $t>0$ and $\theta(t)=0$ for $t<0$, $k=\sqrt{\kappa_b}\omega_{jj'}/c$, $\hbar\omega_{jj'}=E_{1s}(N+1)+E_{jj'}$, and $E_{1s}(N+1)$ is the excitation energy of a free 1s exciton with $q_x=q_y=0$ in an $(N+1)$-monolayer quantum well.1) Since the sum in Eq. (15) includes the squares of the Fourier components $F_\alpha(q_\alpha)$, this sum is nonzero for states with either odd or even indices j and j'. Note that the imaginary unit in the square brackets under the summation sign is compensated by the imaginariness of the square root $\sqrt{q^2-k^2}$ for $k>q$, whereas, when $k<q$, the quantity $P(\sqrt{k^2-q^2})$ is real. Estimates show that the contribution to the sum from the $k>q$ region is negligible for typical quantum wells, and this contribution is included in Eq. (15) for generality. In Fig. 3, the dashed curves show the dependence on L_y of the splitting E_x-E_y between the radiative states $|1,1,x\rangle$ and $|1,1,y\rangle$, calculated for three different values of L_x for a 10-monolayer GaAs/AlGaAs quantum well and an 11-monolayer island. It can be seen from Fig. 3 that, in the range of island sizes considered here, the splitting lies within the limits of ± 50 μeV. This is much less than the energy spacing between the levels of the localized states (of the order of meV or more), and this makes it possible to restrict oneself to first-order perturbation theory when calculating the exchange corrections.

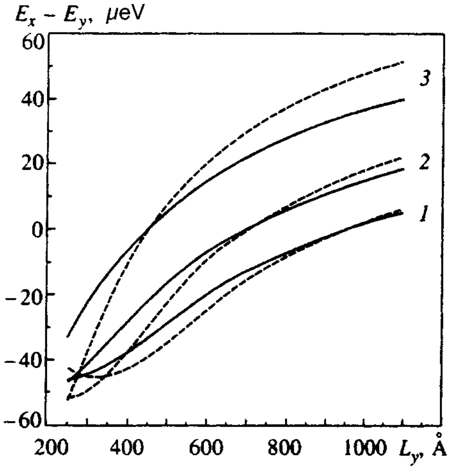


FIG. 3. Splitting of the radiation doublet in a (1,1) exciton, calculated in the model of infinitely high barriers (dashed curves) and in the factorized-envelope approximation (solid curves) for three values of the length $L_x = 950$ Å (*1*), 700 Å (*2*), and 450 Å (*3*).

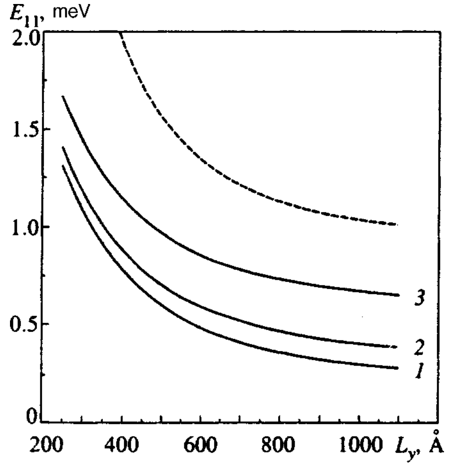


FIG. 4. Energy of the (1,1) ground state of a localized exciton vs length L_y for a fixed value of $L_x = 950$ Å (curve *1*), 700 Å (curve *2*), and 450 Å (curve *3*). The dashed curve is calculated for $L_x = 450$ Å in the model of infinitely high barriers.

5. LOCALIZED EXCITON STATES IN THE FACTORIZED-ENVELOPE APPROXIMATION

The difference $E_{1s}(N) - E_{1s}(N+1) \equiv U$ of the excitation energies of free $1s$ excitons in N- and $(N+1)$-monolayer quantum wells determines the height of the barrier that localizes the exciton in the island. In a GaAs/Al$_{0.3}$Ga$_{0.7}$As structure, the value of U for $N=10$ is 11 meV. In this section, we analyze the effect of the finiteness of barrier U on the exciton levels $E_{jj'}$ and their exchange splitting. For $L_x, L_y > a_B$, the envelope $\Psi_{\text{exc}}(\mathbf{r}_e, \mathbf{r}_h)$ can be approximately represented in the form of Eq. (11), bearing in mind that $f(\boldsymbol{\rho})$, $\varphi_e(z_e)$, and $\varphi_h(z_h)$ are the functions calculated for an $(N+1)$-monolayer well if the coordinates X and Y lie inside the island or for an N-monolayer well in the region outside the island. Then the envelope $F(X,Y)$ satisfies the two-dimensional Schrödinger equation

$$\left[-\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2}\right)-U\theta\left(\frac{L_x}{2}-|X|\right)\theta\left(\frac{L_y}{2}-|Y|\right)\right] \times F(X,Y) = -\varepsilon F(X,Y), \tag{16}$$

where ε is the localization energy of the exciton $(\varepsilon > 0)$, associated with energy $E_{jj'}$ by the relationship $E_{jj'} = U - \varepsilon_{jj'}$. It is assumed that the envelope $F(X,Y)$ and its derivative in the direction perpendicular to the side of the rectangle are continuous on the perimeter of the island. An efficient approximate method of solving such an equation was proposed in Ref. 18. Function $F(X,Y)$ is sought in the form of a product of functions $F_x(X)$ and $F_y(Y)$ that satisfy the system of equations

$$\begin{gathered}\left[-\frac{\hbar^2}{2M}\frac{d^2}{dX^2}-UP_y\theta\left(\frac{L_x}{2}-|X|\right)\right]F_x(X) = -\varepsilon_x F_x(X),\\ \left[-\frac{\hbar^2}{2M}\frac{d^2}{dY^2}-UP_x\theta\left(\frac{L_y}{2}-|Y|\right)\right]F_y(Y) = -\varepsilon_y F_y(Y),\end{gathered} \tag{17}$$

coupled with each other via the quantities

$$P_x = \int_{-L_x/2}^{L_x/2} F_x^2(X)\,dX, \quad P_y = \int_{-L_y/2}^{L_y/2} F_y^2(Y)\,dY.$$

These quantities differ from unity to the extent that functions $F_x(X)$ and $F_y(Y)$ penetrate under the potential barriers. The localization energy of the exciton is expressed in terms of auxiliary energies ε_x and ε_y in the form $\varepsilon = \varepsilon_x + \varepsilon_y - UP_xP_y$. As before, the localized states are characterized by the pair of indices j and j', where $j-1$ and $j'-1$ determine the number of zeroes of functions $F_x(X)$ and $F_y(Y)$, respectively. In Fig. 4, the solid curves show the localization energy of the ground state of an exciton bound on an island in a GaAs/AlGaAs(001) quantum well whose width fluctuates from 10 to 11 monolayers. For comparison, the same figure also shows the $E_{11}(L_y)$ dependence calculated for $L_x = 450$ Å in the model of infinitely high barriers $U \to \infty$ (dashed curve). In this case we have $P_x = P_y = 1$, $\varepsilon \to \infty$, while $U - \varepsilon$ remains a finite quantity. A comparison of the dashed curve with solid curve *3* shows that the model of infinitely high barriers rather strongly exceeds the energy values of the localized excitons. The discrepancy between the results of the two calculations increases even further for excited levels $E_{jj'}$ with $j + j' > 2$. On the other hand, replacing P_x and P_y with unity in the system of Eqs. (17), as a consequence of which this system splits into two uncoupled one-dimensional Schrödinger equations, produces results that virtually coincide with the calculation in the approximation of factorized envelopes.

6. DISCUSSION OF THE RESULTS

The solid curves in Fig. 3 were calculated from Eq. (15) in the approximation of a factorized envelope $F(X,Y)$, i.e., with functions that satisfy Eqs. (17) substituted into Eq. (15). It can be seen that, unlike energy E_{11}, the splitting $E_x - E_y$ is not very sensitive to the method of calculation. Figure 5 shows how the energy $E_{jj'}$ and splitting $E_{jj',x} - E_{jj',y}$ depend on the length of one of the sides of the rectangle for a

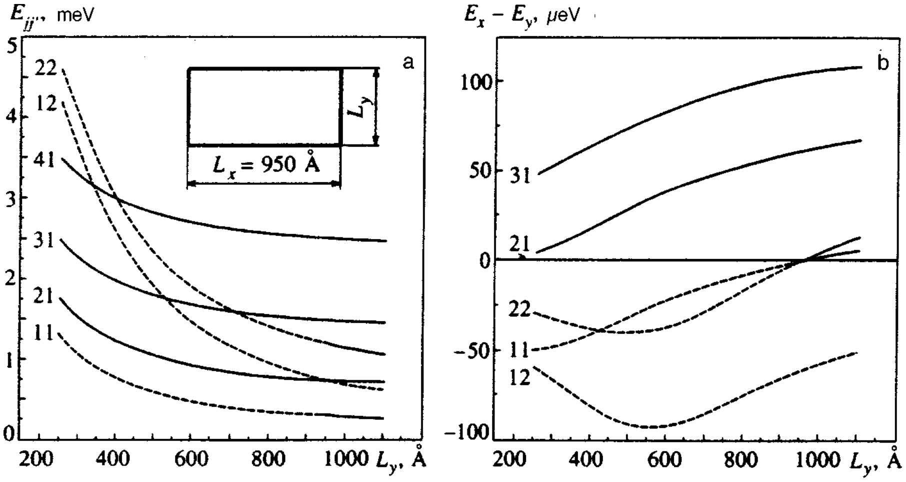


FIG. 5. Energy levels $E_{jj'}$ of a localized exciton (a) and splitting $E_x - E_y$ between localized states $|j,j',x\rangle$ and $|j,j',y\rangle$ (b) vs length L_y for a fixed value of $L_x = 950$ Å (see inset) in a GaAs/Al$_{0.3}$Ga$_{0.7}$As quantum well 28 Å wide ($N = 10$). The curves show pairs of indexes jj' designating the exciton states. The solid and dashed sections of the curves correspond to positive and negative splitting.

fixed length $L_x = 950$ Å of the other side. The sequence of signs of the splitting observed in Ref. 6 is reproduced in the region $L_y = 420-480$ Å.

According to Eq. (15), the sign of the splitting $E_x - E_y \equiv \Delta E_x - \Delta E_y$ is mainly determined by the sign of the difference of the mean squares $\langle q_x^2\rangle$ and $\langle q_y^2\rangle$. Using Eq. (17), it can be shown that $\langle q_x^2\rangle - \langle q_y^2\rangle = (\varepsilon_y - \varepsilon_x)\cdot 2M/\hbar^2$. At $j = j'$ the splitting of the $E_{jj',x}$ and $E_{jj',y}$ levels is evidently negative for $L_y < L_x$ and positive for $L_y > L_x$, and this agrees with curves 11 and 22 in Fig. 5. As the quantum number j increases, the dispersion of $\langle q_\nu^2\rangle$ increases, while the energy ε_ν decreases correspondingly. Therefore, for close-lying L_x and L_y, the signs of the differences $E_{jj',x} - E_{jj',y}$ and $j - j'$ coincide.

The short-range exchange interaction also splits the states $|jj',x\rangle$ and $|jj',y\rangle$. In calculating the short-range contribution, the island-induced admixture of the ±1/2 states of the light hole into the states of the heavy hole with momentum projection ±3/2 must be taken into account. Estimates show that this contribution to $E_x - E_y$ can be neglected by comparison with the analogous contribution of the long-range exchange interaction to the extent that the width of the quantum well is small in comparison with the size of the island in the interface plane.

As already pointed out, in consequence of the rectangular symmetry of the localizing potential, states $|j,j',\nu\rangle$ with even j or j' do not interact with photons propagating along the z axis. Optical transitions to the excited states of a localized exciton with odd j,j' are allowed, but the probabilities $w_{jj',\nu}$ of such transitions are a factor of about $(jj')^2$ less than with excitation of the ground state with $j = j' = 1$. The observation in the photoluminescence excitation spectrum[6] of a set of doublets whose components are polarized along the [110] and [110] axes can be explained by assuming that the localizing anisotropic islands extend along one of these axes, but that its shape is not quite invariant to reflection in the (110) and (110) planes. In other words, it can be assumed when the exchange splitting is calculated that these symmetry elements are present in the system, and, for definiteness, a rectangular shape can be chosen for the islands, whereas, when the transition probabilities $w_{jj',\nu}$ are calculated, distortion of the shape of the island can be taken into account, causing the ground state $|1,1,\nu\rangle$ to be mixed into the excited states $|j,j',\nu\rangle$. For all the excited levels, the probabilities $w_{jj',\nu}$ will then be nonzero but small by comparison with $w_{11,\nu}$. It can be assumed as an alternative that the lines observed in the photoluminescence excitation spectrum correspond to the ground states of excitons that are localized on different islands located at a small distance from one another, so that tunneling transitions from one island to another, induced by acoustic phonons, are possible. In this case, alternation of the signs of the splitting $E_{1\bar{1}0} - E_{110}$ would mean that some of the islands are oriented along the [110] axis, while others are oriented along the [110] axis. This conflicts with Fig. 1 from Ref. 6, which shows a scanning electron micrograph of the GaAs surface.

The theory developed here can thus be used to explain the order of magnitude and sign alternation of the splitting of the ground and excited states of an exciton localized on an anisotropic island in a structure with a quantum well, observed in Ref. 6. We assume that the splitting of radiative exciton states detected in GaAs/AlGaAs quantum wells in Ref. 19 is also associated with anisotropy of the localizing islands. In the structures with quantum wells studied by Heller and Bockelmann,[20] optical orientation of localized excitons was observed only when an external magnetic field was applied; it is natural to explain this on the assumption that a radiation doublet was split. In conclusion, note that the character of the orientation of anisotropic islands must strongly depend on the choice of heterojunction and the technology used to grow the structure. In particular, it is possible for the islands to be distributed randomly over direction so that the structure is isotropic in the middle in the plane of the interface. The overwhelming majority of islands are nevertheless anisotropic, and this local anisotropy, which causes splitting of the radiative states of localized excitons, can be effectively studied by the method of optical orientation and by aligning the excitons in a magnetic field.

This work was supported by the Russian Fund for Fundamental Research (Grant 95-02-06038), INTAS (Grant 93-3657-Ext), and also by the Volkswagen Fund.

[*] E-mail: ivchenko@coherent.ioffe.rssi.ru

[1] Within the limits of accuracy of the effective-mass method, the frequency $\omega_{jj'}$ in the expression for k and the quantity $E_g/\hbar$ in Eq. (15) can be replaced with the resonant excitation frequency of a localized exciton.

---

[1] T. Takagahara, J. Lumin. **44**, 347 (1989).

[2] H. Kalt, J. Collett, S. D. Baranovskii, R. Saleh, P. Thomas, Le Si Dang, and J. Cibert, Phys. Rev. B **45**, 4253 (1992).

[3] L. E. Golub, E. L. Ivchenko, and A. A. Kiselev, J. Opt. Soc. Am. B **13**, 1199 (1996).

[4] G. Bastard, C. Delalande, M. H. Meynadier, P. M. Frijlink, and M. Voos, Phys. Rev. B **29**, 7042 (1984).

[5] L. E. Golub and A. A. Kiselev, in *Proceedings of the Twenty-Third International Symposium on Compound Semiconductors*, St. Petersburg, 1996, Inst. Phys. Conf. Ser. No. 155, chap. 9, p. 687; L. E. Golub, Fiz. Tverd. Tela **39**, 1871 (1997) [Phys. Solid State **39**, 1673 (1997)].

[6] D. Gammon, E. S. Snow, B. V. Shanabrook, D. S. Katzer, and D. Park, Phys. Rev. Lett. **76**, 3005 (1996).

[7] S. V. Goupalov, E. L. Ivchenko, and A. V. Kavokin, in *Proceedings of the International Symposium on Nanostructures: Physics and Technology*, St. Petersburg, 1996, p. 322; in *Abstract Workbook of the Ninth International Conference on Superlattices, Microstructures, and Microdevices*, Liège, Belgium, 1996, p. MoPPT-14.

[8] G. E. Pikus and G. L. Bir, Zh. Éksp. Teor. Fiz. **60**, 195 (1971); **62**, 324 (1972) [Sov. Phys. JETP **33**, 108 (1971); **35**, 174 (1972)].

[9] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Nauka, Moscow, 1972; Wiley, New York, 1975).

[10] V. A. Kiselev, B. S. Razbirin, and I. N. Uraltsev, Phys. Status Solidi B **72**, 161 (1975).

[11] E. M. Gamarts, E. L. Ivchenko, G. E. Pikus, B. S. Razbirin, and A. N. Starukhin, Fiz. Tverd. Tela (Leningrad) **22**, 3620 (1980) [Sov. Phys. Solid State **22**, 2119 (1980)].

[12] E. M. Gamarts, E. L. Ivchenko, G. E. Pikus, B. S. Razbirin, V. I. Safarov, and A. N. Starukhin, Fiz. Tverd. Tela (Leningrad) **24**, 2325 (1982) [Sov. Phys. Solid State **24**, 1320 (1982)].

[13] V. M. Agranovich and V. G. Ginzburg, *Crystal Optics with Spatial Dispersion, and Excitons* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1984)

[14] S. Jorda, U. Rössler, and D. Broido, Phys. Rev. B **48**, 1669 (1993).

[15] L. C. Andreani and F. Bassani, Phys. Rev. B **41**, 7536 (1990).

[16] E. L. Ivchenko, Fiz. Tverd. Tela (Leningrad) **33**, 2388 (1991) [Sov. Phys. Solid State **33**, 1344 (1991)].

[17] F. Tassone, F. Bassani, and L. C. Andreani, Phys. Rev. B **45**, 6023 (1992).

[18] G. Bastard and J. Y. Marzin, Solid State Commun. **91**, 39 (1994).

[19] E. Blackwood, M. J. Snelling, R. T. Harley, S. R. Andrews, and C. B. T. Foxon, Phys. Rev. B **50**, 14 246 (1994).

[20] W. Heller and U. Bockelmann, Phys. Rev. B **55**, 4871 (1997).

Translated by W. J. Manthey

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Electronic structure and superconductivity of κ-(BEDT-TTF)$_2$X salts

V. A. Ivanov,*) E. A. Ugolkova, and M. E. Zhuravlev

N. S. Kurnakov Institute of General and Inorganic Chemistry, Russian Academy of Sciences, 117907 Moscow, Russia

(Submitted 30 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 715–733 (February 1998)

We have analyzed the electronic structure and superconducting properties of layered crystals based on BEDT-TTF salts (hereafter denoted as ET), in which intramolecular interaction among electrons is important. For the case of realistic κ-packing of ET-molecules in a plane of ET$_2$ dimers we have calculated the electron density of states. Using the calculated electronic structure, we have analyzed the symmetry of the anisotropic superconducting pairing. The critical value of the effective attraction between electrons for formation of a bound pair in an empty lattice has been estimated. The relation between the nodes of the d-type order parameter on the anisotropic Fermi surface and superconducting properties of the condensate is discussed in the Bardeen–Cooper–Schrieffer (BCS) model. The results are in agreement with the known band parameters of the normal phase and measurements of the magnetic field penetration depth, heat capacity as a function of temperature in the low-temperature range, and NMR in κ-ET$_2$X superconducting salts. © *1998 American Institute of Physics.* [S1063-7761(98)02602-X]

1. INTRODUCTION

The progress in studies of conductivity[1] and superconductivity[2] of organic materials made since their discovery is a clear indication of their importance for both fundamental science and applications. After the synthesis of the bis(ethylendithio)tetrathiafulvalene (BEDT-TTF, in what follows ET) and the discovery of superconductivity in the β-ET$_2$I$_3$ salt at normal pressures,[3] the κ-ET$_2$X compounds[4] have attracted a lot of attention on account of the diversity of their properties. Although their electronic and crystal structure are quite similar and they all have half a hole per ET molecule, we can find among κ-ET$_2$X compounds semiconductors, such as those with X=Cu[N(CN)$_2$]Cl and d_8-Cu[N(CN)$_2$]Br, and both normal and superconducting metals, for example, so-called 10 K-class superconductors with $T_c \simeq 10$ K under ambient pressure, namely X=Cu[N(CN)$_2$]Br, Cu[N(CN)$_2$]CN, and Cu(NCS)$_2$. Organic superconductors of the κ-ET$_2$X family have the highest superconducting transition temperatures among organic superconductors, up to $T_c \simeq 13$ K. At the same time, the κ-ET$_2$X compounds were the first organic metals whose Fermi-surface topology was determined using Shubnikov oscillations.[5,6] Comprehensive experimental studies of recent years[7–9] have yielded enough information to draw a picture of κ-ET$_2$X properties and develop a common approach to them.

In the conducting plane ET molecules are bound in donor ET$_2$ dimers located at lattice sites, and neighboring dimers are aligned perpendicularly to one another (Fig. 1). This configuration is called κ-ordering of ET molecules. Layers of singly charged ET$_2$ cation radicals alternate with polymer layers of linear and multiply bent X anions. The separation between layers is 15 Å, which equals the lattice constant of bismuth-based high-T_c superconductors of the 2-2-1-2 group and is slightly larger than the layer separation in 1-2-3 superconductors (12 Å). ET$_2$ dimers are located at sites of a plane lattice composed of equilateral triangles whose unit cell is a rectangle with sides a and $\sqrt{3}\,a$ containing two ET$_2$ dimers (in what follows we assume that the lattice constant is $a=1$). Quantum-chemical calculations yield a higher electron density at sulphur and carbon atoms in the plane central fragment C$_2$S$_4$ of the ET molecule. The distance between two plane molecules in each dimer is 3.2 Å, whereas the mean distance between geometrical centers of neighboring ET$_2$ dimers is about 8 Å. The preferential localization of electron density peaks in the central fragments of ET molecules and the small overlap between the wave functions of the molecules are responsible for the high energy U_{ET} of intramolecular Coulomb repulsion between electrons and the narrow energy bandwidths in κ-ET$_2$X materials. Note that at present κ-ET$_2$X organic salts are materials with stronger electron correlations than those in inorganic materials, such as HTSCs and transitional metal oxides. Recently the group of layered organic materials has been supplemented with two novel superconductors, namely ET$_4$X with halogen-mercury layers and the organic metal ET$_8$X$'$(CH$_6$H$_5$Cl)$_2$.[10]

The importance of electron–electron interaction in layered organic salts based on the ET molecule was indicated by Bulaevskii[11] and Toyota *et al.*[12] Kino and Fukuyama[13] analyzed magnetic properties of κ-ET$_2$X salts, α-ET$_2$I$_3$, and ET$_2$MHg(SCN)$_4$, taking into account these interactions in the Hartree–Fock approximation.

Strong electron correlations suppress charge density fluctuations near the top of the band and lead to a weaker electron–phonon interaction. At the same time, the C–S and C=C phonon modes do not contribute to the superconducting isotopic shift ΔT_c, as follows from measurements of T_c in κ-ET$_2$X superconducting salts with isotope ^{34}S and

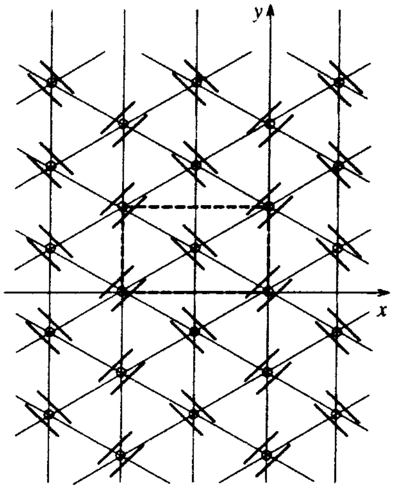


FIG. 1. Alignment of ET_2 dimers at sites of the triangular lattice. The ratio between the sides of the unit cell shown by dashed lines is $1:\sqrt{3}$.

$^{13}C={}^{13}C$ in the central fragment of the ET molecule,[8] nor do they affect normal properties of salts. Thus the role of central vibronic modes and electron–phonon interaction cannot be crucial for normal and superconducting properties of κ-ET_2X salts. Therefore we ignore such modes in our analysis.

Unlike the Hartree–Fock approximation, for generalized Hubbard–Okubo operators[14] the diagram technique allows one to take adequately into account strong electron correlations. This technique describes an insulator–metal transition and indicates that all κ-ET_2X salts that are known at present are close to the phase-transition boundary. The latter is determined by the ratio between the amplitudes for the interdimer and intradimer tunneling of electrons.

The analysis of the electronic structure of κ-ET_2X salts and their superconducting properties described in the paper is based on the energies characterizing these materials, irrespective of the superconducting pairing mechanism: $U_{ET}\geqslant 1$ eV (intramolecular Coulomb repulsion between electrons), $t_0\simeq 0.2$ eV (the integral of electron hopping between ET molecules within ET_2 dimer), $t_{1,2,3}\simeq 0.1$ eV (the integrals of electron hopping between molecules of neighboring dimers). To simplify our calculations we assume that $U_{ET}/t_{1,2,3}\gg 1$. Since there are three electrons for each ET_2^+ dimer, i.e., each hole, $ET_2^+=(ET_aET_b)^+$, the hole representation is preferable in this case. In the effective dimer Hamiltonian approximation[13–16] each dimer is described as a site with two degenerate orbitals, a and b. In other words, κ-ET_2X metals can be described by a doubly degenerate Hubbard model with one carrier per lattice site (dimer) and with two sites (dimers) in one unit cell of a triangular lattice. The electron dispersion laws have been derived in the tight-binding approximation for correlated electrons. The densities of electron states for both normal and superconducting phases have been calculated. Anisotropic superconducting pairing of various symmetry modes has been investigated, and the critical value of the effective attraction between electrons needed for formation of Cooper pairs has been estimated. Strong electron correlations manifest themselves through renormalization of both electron transfer integrals and chemical potential due to correlation factors. It turns out that nodes of the order parameter on the anisotropic Fermi "surface" control such parameters of κ-ET_2X salts as nonactivation electron specific heat, NMR relaxation rate, and penetration depth of a weak magnetic field. The carrier scattering due to nonmagnetic impurities has different effects on coupling constants for symmetry modes with different parities[17] and is inessential for nonspherical Fermi surfaces.[18–21] In this paper we limit our consideration to the case of pure κ-ET_2X crystals without impurities.

2. ELECTRON DENSITY OF STATES IN THE NORMAL PHASE

In κ-ET_2X compounds internal degrees of freedom in the ET_2 plane are important, namely the dimerization of ET molecules and the presence of two ET-dimers in one unit cell. In κ-ET_2X salts with different X^- anions the weak asymmetry of molecule positions in the dimer leads to a difference between the interdimer integrals $t_{1,2,3}$ responsible for intermolecular hopping.[22] Here t_2 is the azimuthal integral of electron hoppings along the b-axis of the crystal lattice,[22] which coincides in this paper with the y-axis. The following electron dispersion functions in the normal (nonsuperconducting) phase for a triangular dimer lattice with different interdimer integrals (the azimuthal integral such as t_2 and nonazimuthal integrals such as t_1 and t_3) in the tight-binding approximation were calculated[16]:

$$\varepsilon_p^{\pm}=t_2\cos p_y\pm\sqrt{t_1^2+t_3^2+2t_1t_3\cos(\sqrt{3}p_x)}\cos\frac{p_y}{2}, \quad (1)$$

where the quasimomenta belong to the first Brillouin zone, $|p_y|\leqslant\pi$, $|\sqrt{3}p_x|\leqslant\pi$. The splitting of the one-particle spectrum (1) is caused by the presence of two dimers in a unit cell. The dimerization of ET molecules leads to further splitting of the spectrum so that each pair of one-particle energies in Eq. (1) generates the upper anti-binding and lower binding bands in accordance with the formula $E_p=\pm t_0\pm\varepsilon_p^{\pm}$, where t_0 is the intradimer hopping integral. Note the two-sheet nature of the two-dimensional Fermi "surface" (Fermi line) due to the branches of spectrum (1). The ε^+ branch defines open sections of the Fermi "surface," and ε^- its closed section. The point $\Gamma(p_x=p_y=0)$ in the Brillouin zone is a saddle point on the dispersion surface $\varepsilon^-(p_x,p_y)$. In the range of small quasimomenta $\mathbf{p}$ the sections ε^+ degenerate to elliptic arcs, and ε^- to a pair of hyperbolic branches.[16] Without taking into account electron correlations, the total carrier density of states (per spin and per unit cell) has the form

$$\rho(\varepsilon)=\sum_{p_x,p_y}[\delta(\varepsilon-\varepsilon^+/t_2)+\delta(\varepsilon-\varepsilon^-/t_2)],$$

$$\rho(\varepsilon)=\frac{1}{\pi^2}\int_{-1}^{1}\frac{dy|\varepsilon-y|}{\sqrt{(1-y)(y_1-y)(y-y_2)(y-y_3)(y-y_4)(y+1)}} \quad (2)$$

with the energy ε expressed in the dimensionless form using the azimuthal hopping integral t_2. In Eq. (2) $y=\cos p_y$ and

$$y_{1,2}=\varepsilon+\frac{\Delta^2}{4}\pm\Delta\sqrt{\frac{1}{2}\left(\varepsilon+1+\frac{\Delta^2}{8}\right)},$$
$$y_{3,4}=\varepsilon+\frac{\delta^2}{4}\pm\delta\sqrt{\frac{1}{2}\left(\varepsilon+1+\frac{\delta^2}{8}\right)},\tag{3}$$

where

$$\Delta=\frac{t_1+t_3}{t_2},\quad \delta=\frac{|t_1-t_3|}{t_2}.$$

The density of states (2) cannot be expressed analytically, but its singular points can be determined by detecting pairwise equality of the roots (3) included in the denominator of the integrand in Eq. (2). It follows from Eq. (2) that the normal electron density of states has split logarithmic peaks at $\varepsilon=-1-\delta^2/8$ and $\varepsilon=1-\Delta$ around the point $\varepsilon=-1$.

In the limiting case of equal nonazimuthal hopping integrals ($\delta=0$, $t_3=t_1$) but $t_2\neq t_{1,3}$ in κ-ET$_2$X salts, one-particle energies (1) are reduced to

$$\varepsilon_p^{\pm}=t_2\cos p_y\pm 2t_1\cos\frac{p_y}{2}\cos\frac{\sqrt{3}p_x}{2}.\tag{4}$$

Then the total density of states is expressed, in accordance with Eq. (2), as

$$\rho(\varepsilon)=\frac{1}{\pi^2}\int_{-1}^{1}\frac{dy}{\sqrt{(1-y)(y_1-y)(y-y_2)(y+1)}},\tag{5}$$

where $y_{1,2}$ are defined by Eq. (3). In the energy band $-(1+\Delta)/2<\varepsilon<1-\Delta$ the density of states can be expressed explicitly as

$$\rho(\varepsilon)=\frac{\sqrt{2}}{\pi^2 q\sqrt{\Delta\sqrt{2+2\varepsilon+\Delta^2/4}}}K\left(\frac{1}{q}\right),$$

$$q=\sqrt{\frac{\left(1-\varepsilon-\frac{\Delta^2}{2}+\frac{\Delta}{2}\sqrt{2+2\varepsilon+\frac{\Delta^2}{4}}\right)\left(1+\varepsilon+\frac{\Delta^2}{2}+\frac{\Delta}{2}\sqrt{2+2\varepsilon+\frac{\Delta^2}{4}}\right)}{2\Delta\sqrt{2+2\varepsilon+\frac{\Delta^2}{4}}}},\tag{6}$$

whereas for the energies $1-\Delta<\varepsilon<1+\Delta$ we obtain

$$\rho(\varepsilon)=\frac{\sqrt{2}}{\pi^2\sqrt{\Delta\sqrt{2+2\varepsilon+\Delta^2/4}}}K(q),\tag{7}$$

where $K(x)$ is the complete elliptic integral of the first kind. As follows from Eqs. (6) and (7), the density of states diverges logarithmically at the energies $\varepsilon=-1$ and $\varepsilon=1-\Delta$.

In the limiting case of fully isotropic interdimer hoppings ($t_3=t_1=t_2$) Eqs. (6) and (7) coincide with the electron density of states given in Refs. 14 and 15. Note that the density of states used in Ref. 23 was determined for a different energy band, so the triangular lattice was considered as a square one with electron hopping to second nearest neighbors in the diagonal direction.

In the specific case of one dimensionality, $\Delta=0$ and the resulting density of states transforms to the well-known form

$$\rho(\varepsilon)=\frac{2}{\pi^2\sqrt{(1-\varepsilon)(1+\varepsilon)}}K(0).$$

It does not contain a logarithmic peak, but square-root Van Hove singularities occur on the boundaries of the allowed energy bands ($\varepsilon=-1$, $\varepsilon=1$) owing to energy factors in front of the elliptic integrals in Eqs. (6) and (7).

3. ANISOTROPIC SUPERCONDUCTING PAIRING

In the study of superconductivity in κ-ET$_2$X salts,[14] only the effective attraction between electrons of nearest neighbors on a square lattice was taken into account, whereas the dispersion equations were derived with due account of diagonal electron hoppings to the next nearest neighbors. In the reference system defined in this paper, the effective attraction between carriers at nearest sites of the regular triangular lattice with the one-point basis is determined by the equation

$$\begin{aligned}V(\mathbf{p}-\mathbf{p}')=2\Bigg[&V_1\cos(p_y-p_y')+V_2\\&\times\cos\frac{\sqrt{3}(p_x-p_x')+p_y-p_y'}{2}+V_3\\&\times\cos\frac{\sqrt{3}(p_x-p_x')-(p_y-p_y')}{2}\Bigg].\end{aligned}\tag{8}$$

Equation (8) for the effective attraction between electrons does not change the symmetry of the dispersion relations for elementary excitations, irrespective of the superconducting pairing mechanism. Its expansion in basis functions $\{\eta_i(\mathbf{p})\}$ of the irreducible representation of the triangular lattice point group in the isotropic case ($V_{1,2,3}=V$) has the form

$$V(\mathbf{p}-\mathbf{p}')=2V\sum_{i=1}^{6}\eta_i(\mathbf{p})\eta_i(\mathbf{p}'),\tag{9}$$

where

$$\eta_1(\mathbf{p})=\frac{1}{\sqrt{3}}\left(\cos p_y+2\cos\frac{p_y}{2}\cos\frac{\sqrt{3}p_x}{2}\right),$$

$$\eta_2(\mathbf{p})=\frac{2}{\sqrt{6}}\left(\cos p_y-\cos\frac{p_y}{2}\cos\frac{\sqrt{3}p_x}{2}\right),$$

$$\eta_3(\mathbf{p})=\sqrt{2}\,\sin\frac{p_y}{2}\sin\frac{\sqrt{3}p_x}{2},$$

$$\eta_4(\mathbf{p})=\frac{1}{\sqrt{3}}\left(\sin p_y+2\sin\frac{p_y}{2}\cos\frac{\sqrt{3}p_x}{2}\right),$$

$$\eta_5(\mathbf{p})=\frac{2}{\sqrt{6}}\left(\sin p_y-\sin\frac{p_y}{2}\cos\frac{\sqrt{3}p_x}{2}\right),$$

$$\eta_6(\mathbf{p})=\sqrt{2}\,\cos\frac{p_y}{2}\sin\frac{\sqrt{3}p_x}{2}. \tag{10}$$

Here the basis functions $\eta_1(\mathbf{p})$, $\eta_2(\mathbf{p})$, and $\eta_3(\mathbf{p})$ describe the isotropic singlet s^*-pairing, $d_{x^2-y^2}$-pairing, and d_{xy}-pairing, respectively, in the chosen reference frame (see Sec. 1), and the basis functions $\eta_{4,5,6}(\mathbf{p})$ are linear combinations of basis functions for the triplet p-pairing. Let us consider the general equation of the standard BCS theory for the superconducting order parameter $\Delta(\mathbf{p})$ at finite temperatures:

$$\Delta(\mathbf{p})=\sum_{p',\alpha=\pm}V(\mathbf{p}-\mathbf{p}')\Delta(\mathbf{p}')\times\frac{\tanh[\sqrt{(\xi_p^\alpha(\mathbf{p}'))^2+\Delta^2(\mathbf{p}')}/2T]}{2\sqrt{(\xi_p^\alpha(\mathbf{p}'))^2+\Delta^2(\mathbf{p}')}}, \tag{11}$$

where the one-particle energies of correlated carriers are

$$\xi_p^\alpha=f\varepsilon_p^\alpha-\mu \tag{12}$$

($f=1/4$ is the correlation factor[14–16] for the actual band filling in κ-ET$_2$X salts with one hole per dimer), the energies ε_p are defined by Eq. (4) at $t_1=t_2$, and μ is the chemical potential. Summation is performed over carrier quasimomenta $\mathbf{p}$ in the ET$_2$ plane.

In reality, closed sections of the Fermi surface ($\xi_p^-=0$) are separated from the open sections ($\xi_p^+=0$) by a small gap owing to the weak anisotropy of the interdimer hopping integral ($t_1\neq t_2$) so that transitions of Cooper pairs between the sections of the Fermi surface with different curvatures can be neglected (11). In the general case the two-band model of superconductivity should be considered.[24,25]

In order to determine the type of the Cooper pairing in our model, let us expand the expression for the gap in Eq. (11) in terms of the basis functions $\{\eta_i(\mathbf{p})\}$ in Eq. (10):

$$\Delta(\mathbf{p})=\sum_{i=1}^{6}\Delta_i\eta_i(\mathbf{p}).$$

As a result, the integral equation for the critical temperature T_c deriving from Eq. (11) with due account of expansion (9) is reduced to a system of algebraic equations

$$\Delta_i=2V\sum_{\substack{\mathbf{p},\\ j=1,\alpha=\pm}}^{6}\frac{\tanh(\xi_p^\alpha/2T_c)}{2\xi_p^\alpha}\Delta_j\eta_i(\mathbf{p})\eta_j(\mathbf{p}). \tag{13}$$

In this case, the temperature T_c is determined by the solvability condition for Eq. (13), i.e., the vanishing of the 6×6 determinant:

$$\left|\delta_{ij}-2V\sum_{\mathbf{p},\alpha=\pm}\frac{\tanh(\xi_p^\alpha/2T_c)}{2\xi_p^\alpha}\Delta_j\eta_i(\mathbf{p})\eta_j(\mathbf{p})\right|=0. \tag{14}$$

For the triangular lattice under consideration most of the nondiagonal elements of the matrix defined by Eq. (14) are zero because the integrands are odd with respect to the variables p_x and p_y, therefore the only pairs of coupled functions are $\eta_1(\mathbf{p})$, $\eta_2(\mathbf{p})$ and $\eta_4(\mathbf{p})$, $\eta_5(\mathbf{p})$.

Measurements of the Knight shift[7,8] give evidence in favor of singlet superconductivity in κ-ET$_2$X salts. Anisotropic singlet d- and s^*-pairing types with nodes of the order parameter on the Fermi ''surface,'' as follows from Eq. (14), decompose into the one-dimensional d_{xy}-pairing and mixed $s^*+d_{x^2-y^2}$-pairing with the basis functions $\eta_3(\mathbf{p})$ and $\eta_1(\mathbf{p})+\eta_2(\mathbf{p})$, respectively. In the case when all constants of effective attraction are different, $V_1\neq V_2\neq V_3$ [Eq. (8)], all the types of singlet pairing are intermixed, but separated from the triplet p-wave pairing.

4. SYMMETRY OF THE SUPERCONDUCTING ORDER PARAMETER

Next we must calculate in the secular equation (14) the following matrix elements

$$\sum_{\mathbf{p},\alpha=\pm}\frac{\tanh(\xi_p^\alpha/2T_c)}{2\xi_p^\alpha}\eta_i(\mathbf{p})\eta_j(\mathbf{p})=\int_{-\infty}^{\infty}d\varepsilon\frac{\tanh(\xi_p^\alpha/2T_c)}{2\xi_p^\alpha}F_{ij}(\varepsilon), \tag{15}$$

where

$$\begin{aligned}F_{ij}(\varepsilon)&=\sum_{\mathbf{p}}\delta(\varepsilon-\varepsilon_p^+)\eta_i(\mathbf{p})\eta_j(\mathbf{p})+\sum_{\mathbf{p}}\delta(\varepsilon-\varepsilon_p^-)\eta_i(\mathbf{p})\eta_j(\mathbf{p})\\&=\frac{\sqrt{3}}{4\pi^2}\int_{-\pi}^{\pi}dp_y\int_{-\pi/\sqrt{3}}^{\pi/\sqrt{3}}dp_x\delta\left(\varepsilon-\cos p_y-2\cos\frac{p_y}{2}\cos\frac{\sqrt{3}p_x}{2}\right)\eta_i(\mathbf{p})\eta_j(\mathbf{p})\\&+\frac{\sqrt{3}}{4\pi^2}\int_{-\pi}^{\pi}dp_y\int_{-\pi/\sqrt{3}}^{\pi/\sqrt{3}}dp_x\delta\left(\varepsilon-\cos p_y+2\cos\frac{p_y}{2}\cos\frac{\sqrt{3}p_x}{2}\right)\eta_i(\mathbf{p})\eta_j(\mathbf{p}).\end{aligned} \tag{16}$$

The resulting functions $F_{ij}(t)$ defined by Eq. (16) with variable $t=\sqrt{3+2\varepsilon}$ have the following forms for different types of superconducting singlet pairing:

$$F_{11}(t)=\frac{1}{3\pi^2}\left\{4I_2-2(1+t^2)I_1+\left[\frac{(t^2-1)^2}{4}-2-t^2\right]I_0\right\},$$

$$F_{22}(t)=\frac{1}{6\pi^2}\left\{I_2+(t^2-5)I_1+\left[\frac{(t^2-1)^2}{4}+1-t^2\right]I_0\right\},$$

$$F_{33}(t)=\frac{1}{2\pi^2}\left\{I_2-\frac{(t^2-1)^2}{2}I_{-1}-(3+t^2)I_1+\left[\frac{(t^2-1)^2}{4}-1+t^2\right]I_0\right\},$$

$$F_{12}(t)=\frac{\sqrt{2}}{3\pi^2}\left\{I_2-\frac{t^2-11}{4}I_1-\left[\frac{(t^2-1)^2}{8}+2-\frac{t^2}{2}\right]I_0\right\}. \quad (17)$$

Here we use the functions

$$I_0=\int_{-1}^{1}\frac{dy}{\sqrt{G(y)}},\quad I_1=\int_{-1}^{1}\frac{y\,dy}{\sqrt{G(y)}},$$

$$I_{-1}=\int_{-1}^{1}\frac{dy}{(y+1)\sqrt{G(y)}},$$

$$I_2=\int_{-1}^{1}\frac{(y+1)^2dy}{\sqrt{G(y)}}=\frac{t^2+3}{2}(I_0+I_1)-\left(\frac{t^2-1}{2}\right)^2I_{-1}+\frac{\sqrt{G(y)}}{y+1}.$$

$$G(y)=(y+1)\left(y-\frac{t^2-1-2t}{2}\right)\left(\frac{t^2-1+2t}{2}-y\right)(1-y), \quad (18)$$

which can be expressed in terms of elliptic integrals. Specifically, in the case of dimensionless energies in the range $-3/2<\varepsilon<-1$, i.e., $0<t<1$, the integrals in Eq. (18) reduce to quadratures:

$$I_0=\frac{4}{\sqrt{(3-t)(1+t)^3}}K(k),$$

$$I_1=\frac{4}{\sqrt{(3-t)(1+t)^3}}\left[-K(k)+\frac{(t-1)^2}{2}\Pi\left(\frac{\pi}{2},\frac{4t}{(1+t)^2},k\right)\right],$$

$$I_{-1}=\frac{4}{(t-1)^2\sqrt{(3-t)(1+t)^3}}\left[\frac{(3-t)(t+1)}{2}E(k)+\frac{(t-1)^2}{2}K(k)\right], \quad (19)$$

where $k^2=16t/(3-t)(1+t)^3$.

In the energy interval $-1<\varepsilon<3$, i.e., $1<t<3$, the integrals (18) are expressed in the following form:

$$I_0=\frac{1}{\sqrt{t}}K(q),$$

$$I_1=\frac{1}{\sqrt{t}}\left[-K(q)+\frac{(t-1)^2}{2}\Pi\left(\frac{\pi}{2},\frac{(3-t)(t+1)}{4},q\right)\right],$$

$$I_{-1}=\frac{4}{(t-1)^2(t+1)^2\sqrt{t}}\left[2tE(q)+\frac{(t-1)^2}{2}K(q)\right] \quad (20)$$

with the modulus $q=1/k$. In Eqs. (19) and (20) the functions $K(k)$, $E(k)$, and $\Pi(\pi/2,n,k)$ denote complete elliptic integrals of the first, second, and third kind, respectively, in the normal Legendre form.

Of all possible symmetry types of superconducting phases described by solutions of Eq. (14), the phase with the highest superconducting transition temperature T_c is realized. The symmetry of the corresponding superconducting order parameter determines the desired form of the anisotropic pairing. The matrix elements (15) of Eq. (14) are transformed to

$$\int_{-\infty}^{\infty}d\varepsilon\,\frac{\tanh(\xi_p/2T_c)}{2\xi_p}F_{ij}(\varepsilon)=\int_{-\omega_c}^{\omega_c}d\varepsilon\,\frac{\tanh(\xi_p/2T_c)}{2\xi_p}F_{ij}(\varepsilon)=\int_{\xi(-\omega_c)}^{\xi(\omega_c)}d\xi\frac{\tanh(\xi/2T_c)}{2\xi f}F_{ij}(\varepsilon(\xi)),$$

where ω_c is the cut-off parameter of the pairing interaction (8) in Eqs. (13) and (14). In the latter equation the singularity point $\xi=0$ corresponds to $\varepsilon=\mu/f$, where the chemical potential $\mu/f=-0.415$ was calculated in Ref. 16 and corresponds to the real filling of the energy bands in κ-ET$_2$X salts.

By applying the logarithmic approximation to the elements of secular equation (14), we find out that for the d_{xy}-wave pairing the transition temperature T_c is determined by a single equation

$$1=\frac{2V}{f}F_{33}\ln\frac{\omega_c f}{2T_c}, \quad (21)$$

so the corresponding coupling constant at $f=1/4$ [cf. Eq. (17)] is

$$\lambda_{xy}=8VF_{33}(1.47). \quad (22)$$

The critical temperature for superconductivity with the order parameter of mixed symmetry $s^*+d_{x^2-y^2}$ is determined by the quadratic equation

$$\begin{vmatrix} 1-8VF_{11}\ln\dfrac{\omega_c}{8T_c} & -8VF_{12}\ln\dfrac{\omega_c}{8T_c} \\ -8VF_{12}\ln\dfrac{\omega_c}{8T_c} & 1-8VF_{22}\ln\dfrac{\omega_c}{8T_c} \end{vmatrix}=0. \quad (23)$$

In the calculation of the coefficients F_{ij} in the matrix elements [Eqs. (17)–(20)] the following numerical values of elliptic integrals are useful: $K(0.99)=3.35$; $E(0.99)=1.03$; $\Pi(\pi/2;0.94;0.99)=24.65$. After substituting them into Eq. (16), we obtain

$$F_{11}(1.47)=\frac{1.60}{\pi^2};\quad F_{22}(1.47)=\frac{0.36}{\pi^2};$$

$$F_{33}(1.47)=\frac{2.09}{\pi^2};\quad F_{12}(1.47)=\frac{(-0.48)}{\pi^2}.$$

Then it follows from Eq. (23) that the coupling constants of the mixed-symmetry $d_{x^2-y^2}+s^*$-pairing are $\lambda_{x^2-y^2+s^*}=8VF_{1,2}$, where $F_1=1.76/\pi^2$ and $F_2=0.20/\pi^2$, i.e, both constants are smaller than the coupling constant in Eq. (21) for the superconducting d_{xy}-wave pairing. It follows that the pairing type with the d_{xy}-symmetry is preferable in our model of superconducting κ-ET$_2$X salts.

For the amplitude of the effective pair interaction $V_0=0.022$ transformed to the dimensionless form using the interdimer hopping integral t, the ratio of the critical temperatures is $T_c^{d_{xy}}/T_c^{s^*+d_{x^2-y^2}}=1.65$. If the cut-off parameter ω_c of the effective attraction V is approximately equal to the hopping integral, i.e., $\omega_c\approx t\simeq 0.1$ eV, in the logarithmic approximation the superconducting transition temperature T_c for the d_{xy}-wave pairing, which is realized in experiment, is 10 K [Eq. (21)].

5. ELECTRON PAIR IN AN EMPTY LATTICE OF ET$_2$ DIMERS

In the previous section we calculated the superconducting pairing constant and showed that anisotropic d-pairing should occur. In the case of scattering of a pair of particles in the Cooper s-channel (with zero orbital quantum number $l=0$), the positive phase shift $\delta_{l=0}(E_F)>0$ due to scattering on the two-dimensional lattice at a nonzero Fermi energy the (the Cooper superconducting instability) entails the phase shift $\delta_{l=0}(0)>0$ in an empty lattice (bound two-particle state).[26,27]

Unlike the case of s-pairing, the formation of a Cooper pair and a bound state of two electrons in the empty lattice are essentially different problems in the case of the d- and p-pairing, and the solutions of these problems should be obtained independently. In the model of superconductivity in κ-ET$_2$X under consideration, Cooper pairs can be formed at intensities of the effective interelectron attraction smaller than that needed for pair formation in the empty lattice. In order to calculate the critical value of the interaction amplitude at which a bound electron pair is formed in an empty triangular lattice, one should start with the Bethe–Salpeter equation for the scattering T-operator, which is expressed as

$$T_{pp'}(E)=V_{pp'}+\frac{\Omega}{(2\pi)^2}\int\int\frac{V_{pp'}T_{p'p''}(E)}{E-2\varepsilon_p^-}dp''_x dp''_y, \tag{24}$$

where ε_p^- is the one-particle energy derived from Eq. (4), and $\Omega=\sqrt{3}$ is the volume of the unit cell of the dimer lattice under consideration (its constants $a=1$, as stated in the Introduction).

In the case of the anisotropic d_{xy}-pairing described in the previous section the pairing interaction is described by the equation

$$V_{pp'}=-2V\sin\frac{p_x}{2}\sin\frac{p_y}{2}\sin\frac{p'_x}{2}\sin\frac{p'_y}{2}.$$

Assuming that the scattering operator has the same symmetry and substituting the expression for it in the form

$$T_{pp'}(E)=C\sin\frac{p_x}{2}\sin\frac{p_y}{2}\sin\frac{p'_x}{2}\sin\frac{p'_y}{2}$$

into Eq. (24), we obtain an equation for the amplitude C of the scattering operator:

$$C=-2V-C\frac{V}{t}\frac{\Omega}{(2\pi)^2}\times\int\int\frac{dy\,dz}{\varepsilon-z}\sqrt{1-\left[\frac{\cos y-z}{2\cos(y/2)}\right]^2}\left(\sin\frac{y}{2}\right)^2, \tag{25}$$

where the integration is performed over the region of x and y where the integrand in Eq. (25) is real. Since we are seeking the condition for formation of a bound pair in an empty lattice, the pair energy ε is slightly lower than the bottom of the band, $\varepsilon_{\min}^-/t=-3/2$. In order to estimate the integral in Eq. (25), we replace the factors in the integrand outside the square root with their values at the maximum of the integrand. Thereafter one can easily obtain the critical value of the interaction amplitude V at which the scattering operator changes its sign and a bound state of two electrons occurs:

$$V^{\mathrm{cr}}\simeq\frac{2\pi}{3}t, \tag{26}$$

where the numerical factor is determined by the triangle lattice symmetry.

It is noteworthy that the critical value of the interaction amplitude given by Eq. (26) for a bound pair in the d_{xy}-channel of our triangle lattice ET$_2$ is smaller than that for a square lattice in the Hubbard model,[28] but is slightly larger than the critical value calculated for the $t-J$ model on a square lattice.[29] Earlier[30–32] it was derived from the Hubbard model in the mean-field approximation by assuming that involvement of second nearest neighbors in a square lattice also reduces the critical threshold attraction needed for s-pairing.

6. DENSITY OF ELECTRON STATES IN THE SUPERCONDUCTING PHASE

The nodes of the superconducting order parameter Δ_d obtained in Sec. 4 on both open and closed sections of the two-sheet Fermi surface manifest themselves in the density of electron states in the superconducting phase and are essential for interpretation of such measured physical parameters as the electron heat capacity, NMR relaxation rate, magnetic field penetration depth in the superconducting state, etc.

The density of states in the superconducting phase for energy $E>0$ is determined by the formula

$$\rho_s^{\pm}(E)=\frac{\sqrt{3}}{4\pi^2}\int_{-\pi}^{\pi}dp_y\int_{-\pi/\sqrt{3}}^{\pi/\sqrt{3}}\delta(E-\sqrt{(\xi_p^{\pm})^2+\Delta_p^2})dp_x, \tag{27}$$

where the order parameter is

$$\Delta_p = \Delta_0 \sin \frac{p_y}{2} \sin \frac{\sqrt{3} p_x}{2}$$

and one-particle energies of correlated carriers are

$$\xi_p^{\pm} = ft \left(\cos p_y \pm 2 \cos \frac{p_y}{2} \cos \frac{\sqrt{3} p_x}{2} \right) - \mu,$$

[cf. Eq. (12)]. In deriving Eq. (27) we have used the fact that the coherence factors $1 \pm \xi_p / \sqrt{\xi_p^2 + \Delta_p^2}$ cancel out under integration because the second terms are odd with respect to the energy ξ_p.

The parameter Δ_p is small near its four nodes on the Fermi surface within the first Brillouin zone around the lines $p_x = 0$ and $p_y = 0$. Let us expand ξ_p and Δ_p in Eq. (26) in powers of their small deviations from their values at the nodes of the order parameter, $\Delta_p = 0$, on the Fermi surface $\mu(p_x, p_y)$. From these calculations we derive the density of electron states $\rho_s^+(E)$ as

$$\rho_s^+(E) = \frac{E}{2\pi \Delta_0 ft \sin^2 \frac{p_y}{2} \left(2 \cos \frac{p_y}{2} + 1 \right)} \tag{28}$$

near the node

$$p_x = 0, \quad p_y = 2 \arccos \left(\sqrt{\frac{3}{4} + \frac{\mu}{2ft}} - \frac{1}{2} \right)$$

on the open section of the Fermi surface ($\xi^+ = 0$). Similarly, around the node

$$p_y = 0, \quad p_x = \frac{2}{\sqrt{3}} \arccos \left(\frac{1}{2} - \frac{\mu}{2ft} \right)$$

on the closed section of the Fermi surface ($\xi^- = 0$) the density of states is expressed as

$$\rho_s^-(E) = \frac{E}{2\pi \Delta_0 ft \sin^2 \frac{\sqrt{3} p_x}{2}}. \tag{29}$$

Inside conventional superconducting gap of the s-wave type the density of electron states is zero. In the case of the anisotropic d_{xy}-wave type order parameter the density of electron states (28), (29) in the superconducting phase is, naturally, gapless and has a linear dependence on energy near the Fermi surface. Note that sections of the Fermi energy with different curvatures (ξ^+ and ξ^-) yield different contributions to these linear function.

By comparing Eqs. (28) and (29) and taking into account the chemical potential given above, $(\mu/t) f = -0.415$,[16] we find that the factor in the linear density of states due to the saddle section of the Fermi surface, $\xi_p^- = 0$ [Eq. (29)] is larger by a factor of three than that due to the open sheet of the Fermi surface, $\xi_p^+ = 0$ [Eq. (28)]. This fact is essential for interpretation of superconducting phase parameters in layered organic materials, such as κ-$\mathrm{ET}_2\mathrm{X}$.

7. CHARACTERISTICS OF THE SUPERCONDUCTING PHASE

Owing to the small coherence length in superconducting κ-$\mathrm{ET}_2\mathrm{X}$ salts, the London penetration depth $\lambda_{\alpha\alpha}$ of the magnetic field along α-direction is expressed in the local electrodynamic approximation[33,34] by the formula

$$\frac{c^2}{4\pi n e^2} \lambda_{\alpha\alpha}^{-2} = \sum_p \left(\frac{\partial \xi_p}{\partial p_\alpha} \right)^2 \left\{ -\frac{\partial N_F(\xi_p)}{\partial \xi_p} - \left[-\frac{\partial N_F(E_p)}{\partial E_p} \right] \right\}, \tag{30}$$

where n is the concentration of superconducting carriers, α is the direction of the superconducting screening current, N_F is the Fermi–Dirac distribution function, ξ_p is the one-particle electron energy in the normal phase [Eq. (4)] with respect to the Fermi energy μ, and $E_p = \sqrt{\xi_p^2 + \Delta_p^2}$. In calculating the penetration depth of the magnetic field perpendicular to the ET_2 layer, one should substitute into Eq. (30) the appropriate dispersion functions ξ_p, i.e., ξ^+ for the calculation of λ_{xx} and ξ^- for λ_{yy}. At zero temperature the second term in Eq. (30) goes to zero, and the derivative of the Fermi–Dirac distribution function becomes a δ-function. As a result, we have

$$\lambda_{xx}^{-2} \propto \frac{3}{4\pi^2} f^2 \{ -I_2^x + (2\mu + 4) I_1^x + (3 - \mu^2) I_0^x \},$$

$$\lambda_{yy}^{-2} \propto \frac{1}{\pi^2} f^2 \left\{ \frac{3 - \mu^2}{4} I_0^y - \frac{\mu}{2} I_1^y + \frac{(\mu + 1)^2}{2} I_{-1}^y - \frac{1}{4} I_2^y \right\}. \tag{31}$$

Here I_0^i, I_{-1}^i, I_1^i, and I_2^i are determined by elliptic integrals

$$I_2^i = \left(\frac{\mu}{ft} + 3 \right) I_1^i + \left(\frac{\mu}{ft} + 3 \right) I_0^i + \left(\frac{\mu}{ft} + 1 \right)^2 I_{-1}^i + \frac{\sqrt{G(y)}}{y + 1},$$

where integration of functions determining I_0^i, I_{-1}^i, I_1^i, and I_2^i is performed from y_2 to μ for λ_{xx} and from μ to y_1 for λ_{yy}, and y_1 and y_2 are defined by Eq. (3).

Electron correlations generate coefficients f and f^2 in Eq. (31) and renormalize the chemical potential μ, influencing $\lambda_{xx,yy}$ via incomplete elliptic integrals.

By calculating I_0^i, I_{-1}^i, I_1^i, and I_2^i in Eq. (31), we obtain the ratio between penetration depths of magnetic fields aligned with the x- and y-axes in the dimer triangular lattice:

$$\frac{\lambda_{xx}}{\lambda_{yy}} \approx 1.1. \tag{32}$$

The density of states in the superconducting state calculated in the previous section allows us to obtain the electron heat capacity as a function of temperature and the spin–lattice relaxation rate due to conduction electrons in the ET_2 plane. Unlike the magnetic field penetration depth, these parameters are determined by the density of states averaged over both portions of the Fermi surface.

The linear energy dependence of the density of states leads to a quadratic temperature dependence of the electron heat capacity per unit cell of the ET_2 dimer layer:

$$C_s=2\sum_{p\alpha}\xi_p^{\alpha}\frac{\partial N_F}{\partial T}=2\int_0^{\infty}\beta E^2\frac{\partial N_F}{\partial T}dE$$

$$=6\beta T^2\int_0^{\infty}\frac{x^2dx}{e^x+1}=9\beta\zeta(3)T^2=10.8\beta T^2. \qquad (33)$$

Here β is the sum of the coefficients in front of the energy E in the superconducting density of states derived from Eqs. (28) and (29), and ζ is the Riemann zeta-function. Note that the main contribution to the heat capacity (33) is due to electrons on the closed saddle-shaped section of the Fermi surface $\xi_p^-=0$. By substituting realistic parameters for κ-ET$_2$X into Eq. (33), namely $t=0.12$ eV, $\Delta=(2.5-3.5)T_c$,[35,36] and $T_c=10$ K, we obtain the electron heat capacity per mole in the superconducting phase in the form $C_m=\alpha T^2$, where $\alpha=N_A\beta k_B^3/2$ varies between 1.59 and 2.23 mJ/K^3·mol, which is in satisfactory agreement with experimental data,[37] namely 2.2 mJ/K^3·mol for the samples of the κ-(ET)$_2$Cu[N(CN)$_2$]Br superconductor with $T_c=11.6$ K, and less than 3.53 mJ/K^3·mol for κ-(ET)$_2$Cu(NCS)$_2$ with $T_c=10$ K.

At low temperatures the spin–lattice relaxation rate $R=1/T_1$ of nuclear magnetic moments of central ^{13}C atoms of the ET molecule due to conduction electrons is determined in the superconducting phase by the equation[33]

$$R_s\approx\sum_{p}\sum_{p'}\frac{E_pE_{p'}+\Delta_p\Delta_{p'}+\xi_p\xi_{p'}}{2E_pE_{p'}}$$

$$\times\frac{N_F(E_p)-N_F(E_{p'})}{\exp[(E_p-E_{p'})/T]-1}\,\delta(E_{p'}-E_p-\nu), \qquad (34)$$

where $E_p=\sqrt{\xi_p^2+\Delta_p^2}$ and ν is the high frequency of the magnetic field. Near the nodes the following equations apply:

$$\Delta_{p'}=\pm\Delta_p\left(1+\frac{\nu}{E_p}\right),\qquad \xi_{p'}^{\alpha}=\pm\xi_p^{\alpha}\left(1+\frac{\nu}{E_p}\right).$$

The integration regions in Eq. (34) determine the sign change in Δ_p and $\Delta_{p'}$, as well as in ξ_p and $\xi_{p'}$. After calculations we obtain an expression for the ratio of the relaxation rates in the superconducting and normal phases:

$$\frac{R_s}{R_n}=\frac{2}{T}\int_0^{w}\frac{\beta^2E^2}{\rho^2(\mu)}\,\frac{\exp(E/T)}{(e^{E/T}+1)(e^{(E+\nu)/T}+1)}dE. \qquad (35)$$

Here the parameter β is the same as in Eq. (33), ρ is the normal density of states on the Fermi level, w is the boundary energy below which the superconducting density of states is linear. Given that the high frequency ν satisfies ~ 10 MHz$\sim 10^{-4}$ K and at low temperatures we have $w\gg T$, we obtain

$$\frac{R_s}{R_n}=\frac{2T^2\beta^2\zeta(2)}{\rho^2(\mu)}\left(1-\frac{\nu}{T}\,\frac{\ln 2}{\zeta(2)}\right). \qquad (36)$$

In the normal phase, in accordance with the Korringa law, $R_n\propto T$, we, therefore have $R_s\propto T^3$.

Electron correlations affect the factor β in Eqs. (33) and (36) through the correlation factor $f=1/4$ and the renormalized chemical potential.

The resulting function on the right-hand side of Eq. (36) is quite different from the exponentially small function $R_s\propto\exp(-\Delta/T)$, which occurs at low temperatures in superconductors with conventional s-wave pairing.

8. DISCUSSION OF RESULTS

Analytically defined branches (1) of the energy dispersion relations in the model of κ-ET$_2$X salts under consideration provide a two-sheet Fermi surface (Fermi lines) with an open section of positive curvature and a closed saddle-shaped section. The Fermi surface anisotropy manifests itself in features of the density of electron states in the normal (Sec. 2) and superconducting (Sec. 6) phases and, naturally, determines the symmetry of the effective attraction between electrons [Eqs. (8) and (9) in Sec. 3]. In the case of singlet pairing in pure crystals, our model leads to superconductivity with the order parameter of the d-wave symmetry:

$$\Delta_p=\Delta_0\sin\frac{p_y}{2}\sin\frac{\sqrt{3}p_x}{2}$$

(Sec. 4) in the reference system defined in Sec. 2. The order parameter has a pair of nodes on each section of the Fermi surface within the first Brillouin zone along the lines $p_x=0$ and $p_y=0$. According to Secs. 4 and 5, the corresponding constant of superconducting pairing is provided by the amplitude of the effective attraction between carriers, $V<V_c$, which is, however, insufficient for formation of bound electron states in the d-channel in an empty lattice. Properties of the energy dispersion relations (1) manifest themselves not only in the normal density of electron states analyzed in Sec. 2, but also in the density of states in the superconducting condensate phase with the anisotropic order parameter calculated in Sec. 4. The superconducting density of electron states calculated in Sec. 6 has no gaps around the nodes on the Fermi surface and is linear in the energy [Eqs. (28), (29)]. As a result, the number of elementary thermal excitations is not exponential in temperature but proportional to a power function. Consequently, in the superconducting phase at low temperatures the electron heat capacity is quadratic in temperature [Eq. (33)] and the spin–lattice relaxation rate $1/T_1$ is cubic in temperature.

The effect of the topology of the anisotropic Fermi surface on the density of electron states and characteristics of the superconducting phase in the case of isotropic s-wave pairing was investigated long ago.[38] At present, the anisotropic superconducting order parameter is usually analyzed assuming an isotropic Fermi surface and quadratic dispersion relations for electrons. In three-dimensional superconductors with anisotropic d-wave pairing the density of electron states near the nodes of the order parameter on a spherical Fermi surface is proportional to the energy squared, as follows from symmetry considerations[39] or calculations.[40]

In the reported work we have studied the effect of the topology of a two-dimensional Fermi surface on properties of the superconducting condensate, with a particular attention to κ-ET$_2$X salts. According to Eqs. (28) and (29), the density of electron states on the saddle-shaped section of the Fermi surface is a factor of three higher than on the open section, which manifests in the superconducting specific heat. Therefore contributions to the heat capacity proportional to T^2 should be smaller in ET-superconductors without saddle-shaped surfaces. Specifically, Fermi surfaces of $\alpha\text{-ET}_2\text{NH}_4\text{Hg(SCN)}_4$ layered superconductors have no saddle-shaped sections and, as follows from the experimental data,[41,42] there is no substantial contribution to their superconducting heat capacity proportional to T^2. Thus, in analyzing properties of $\kappa\text{-ET}_2\text{X}$ salts we should not describe their Fermi surfaces as circles (cylinders), as is often done[43] by the analogy with free electron gas.

Note, however, that, according to Eq. (32), the Fermi surface anisotropy has little effect on the main values of the tensor of the magnetic field penetration depth in the ET_2 plane.

The calculated cubic temperature dependence of the spin relaxation rate due to conduction electrons is in agreement with experimental data concerning nuclear magnetic moments of carbon isotopes, $R_s(^{13}\text{C}) \propto T^3$ in $\kappa\text{-ET}_2\text{X}$ superconductors[44–46] at low temperatures $T \ll T_c$. The Hebel–Slichter peak is missing in these measurements at $T \leqslant T_c$ because of the strong electron correlations only normal averages contribute to the nuclear spin relaxation rate, whereas one-site anomalous averages like $\langle a_\uparrow^{(+)}(t) a_\downarrow^{(+)} \rangle$ vanish, provided that $U_{\text{ET}}/t_{1,2,3} \gg 1$.[47] The contribution of additional relaxation mechanisms, such as impurity relaxation, however, should not be neglected. In this paper, we discuss only the case of superconductors without impurities and neglect damping of elementary excitations. Equation (34) does not apply to the proton spin–lattice relaxation on the edge fragments of ET molecules, where the relaxation due to normal electrons in vortex cores should be taken into account.

Note that a quadratic temperature dependence of the electron heat capacity was detected previously in 1-2-3 high-T_c superconductors.[48] A cubic dependence of the spin–lattice relaxation rate, $R_s \propto T^3$,[49,50] and the absence of the Knight shift were also detected in cuprate high-T_c superconductors. These experimental results were interpreted in terms of superconducting pairing with the order parameter of the $d_{x^2-y^2}$-wave symmetry in CuO_2 layers.[51,52] This pairing type corresponds to the d_{xy}-wave symmetry in the reference system used in our analysis.

The correlation factor $f=(4-3n)/4$, where $n=1$ is the number of holes per dimer, affects the properties of the superconducting condensate, including T_c. In the normal phase this factor alters the energy dispersion relation (1) to (12) for correlated carriers. The resulting four-fold narrowing of the conduction energy band may be responsible for the difference between optical and cyclotron masses[53–55] and for the reduction in the areas of hole orbits derived from quantum magnetic oscillations in the $\kappa\text{-ET}_2\text{Cu[N(CN)}_2\text{]Br}$ salt.[56] It follows from Eq. (12) that the gap between the closed and open portions of the Fermi surface required for the magnetic breakdown should be

$$\xi_p^+ - \xi_p^- = 2f \cos\frac{p_{y0}}{2}\,|t_1 - t_3| \simeq \frac{|t_1 - t_3|}{4} \approx 4\ \text{meV}$$

$(p_{y0} \approx 2\pi/3)$ for a realistic difference between nonazimuthal hopping integrals t_1 and t_3. This value is in agreement with experimental data.[57–60]

Komatsu *et al.*[61] detected in their copper ESR signals a change in the Cu^{2+} concentration on the acceptor layer of the $\kappa\text{-ET}_2\text{Cu}_2\text{(CN)}_3 = \kappa\text{-ET}_2^{1-x}\text{Cu}_{2-x}^{+}\text{Cu}_x^{2+}(\text{CN}^-)_3$ superconductor, i.e., an increase in the concentration of Cu^{2+} paramagnetic ions leads to a reduction in the concentration $1-x$ of holes in the ET_2 donor layer, which leads to an increase in the correlation factor $f=(1+3x)/4$ and a decrease in T_c, in accordance with Eqs. (20) and (21) of Sec. 4. This conclusion is in agreement with experimental data.[61]

The existing pressure technique allows the experimenters to apply stress along certain crystal axes of synthesized $\kappa\text{-ET}_2\text{X}$ samples. In this connection, the proposed analytical formulas describing the electronic structure and properties of $\kappa\text{-ET}_2\text{X}$ salts in terms of the intradimer hopping integral t_0 and interdimer hopping integrals $t_{1,2,3}$ could be useful.

It is noteworthy that Eqs. (33) and (34) yield the following formula relating the nuclear spin relaxation rate (in zero magnetic field) and the electron heat capacity in the superconducting phase of $\kappa\text{-ET}_2\text{X}$:

$$\frac{R_s C_n^2}{R_n C_s^2} = \frac{4\zeta^3(2)}{81\zeta^2(3)} = 0.29, \tag{37}$$

where the normal electron heat capacity is defined as $C_n = 2\zeta(2)\rho(\mu)T$. It is of interest to check Eq. (37) in experiments, since a deviation from Eq. (3) would indicate the existence of additional relaxation mechanisms.

The work was supported by the Russian Fund for Fundamental Research.

*)E-mail: vonavi@ionch.ran.msk.ru

---

[1] H. Akamatsu and H. Inokuchi, J. Chem. Phys. **18**, 810 (1950); H. Akamatsu, H. Inokuchi, and Y. Matsunaga, Nature **173**, 168 (1954).

[2] D. Jerome, A. Mazaud, M. Ribault, and K. Bechgaard, J. Phys. Lett. **41**, L95 (1980); K. Bechgaard, C. S. Jacobsen, K. Mortensen *et al.*, Solid State Commun. **33**, 1119 (1980).

[3] É. B. Yagubskiĭ, I. F. Shchegolev, V. N. Laukhin *et al.*, JETP Lett. **39**, 12 (1984); V. F. Kaminskii, T. G Prokhorova, R. P. Shibaeva *et al.*, JETP Lett. **39**, 17 (1984).

[4] H. Urayama, H. Yamochi, G. Saito *et al.*, Chem. Lett. **1**, 55 (1988).

[5] M. V. Kartsovnik, V. N. Laukhin, V. I. Nizhankovskii, and A. A. Ignat'ev, JETP Lett. **47**, 363 (1988); M. V. Kartsovnik, P. A. Kononovich, V. N. Laukhin, I. F. Shchegolev, JETP Lett. **48**, 541 (1988).

[6] K. Murata, N. Toyota, Y. Honda *et al.*, J. Phys. Soc. Jpn. **57**, 1540 (1988).

[7] D. Jerome, in *Organic Conductors*, ed. by J.-P. Farges, Marcel Dekker (1994), Ch. 10, p. 405.

[8] K. Kanoda, Kotai Buzuri, Solid State Phys. (Tokyo) [in Japanese] **30**, 240 (1995).

[9] T. Ishiguro and K. Yamaji, *Organic Superconductors, Solid State Science* Vol. 88, Springer-Verlag, Tokyo (1990).

[10] R. B. Lyubovskiĭ, S. I. Pesotskiĭ, A. Gilevskii, and R. N. Lyubovskaya, Zh. Éksp. Teor. Fiz. **107**, 1698 (1995) [JETP **80**, 946 (1995)].

[11] L. N. Bulaevskii, Adv. Phys. **37**, 443 (1988).

[12] N. Toyota, E. W. Fenton, T. Sasaki, and M. Tachiki, Solid State Commun. **72**, 859 (1989).

[13] H. Kino and H. Fukuyama, J. Phys. Soc. Jpn. **65**, 2158 (1996).

[14] V. Ivanov and K. Kanoda, Physica C **268**, 205 (1996); Mol. Cryst. Liq. Cryst. **285**, 211 (1996).

[15] V. A. Ivanov, Physica C **271**, 127 (1996).

[16] V. Ivanov, K. Yakushi, and E. Ugolkova, Physica C **275**, 26 (1997).

[17] E. Z. Kuchinskiĭ, M. V. Sadovskiĭ, M. A. Érkabaev, Zh. Éksp. Teor. Fiz. **104**, 3550 (1993) [JETP **77**, 692 (1993)].

[18] R. Fehrenbacher and M. R. Norman, Phys. Rev. B **50**, 3495 (1994).

[19] A. A. Abrikosov, Physica C **214**, 103 (1993).

[20] L. A. Openov, V. F. Elesin, and A. V. Krasheninnikov, Physica C **275**, 53 (1996).

[21] A. I. Posazhennikova and M. V. Sadovskiĭ, JETP Lett. **63**, 358 (1996).

[22] K. Oshima, T. Mori, H. Inokuchi *et al.*, Phys. Rev. B **38**, 938 (1988).

[23] Y. Hasegawa, P. Lederer, T. M. Rice *et al.*, Phys. Rev. Lett. **63**, 907 (1989).

[24] V. A. Moskalenko, M. E. Palistrant, and V. M. Vakalyuk, Usp. Fiz. Nauk **161**, 155 (1991) [Sov. Phys. Usp. **34**, 265 (1991)].

[25] A. A. Gorbatsevich and Yu. V. Kopaev, JETP Lett. **51**, 373 (1990).

[26] K. Miyakc, Prog. Theor. Phys. **69**, 1794 (1983).

[27] M. Randeria, J.-M. Duan, and L.-Y. Shieh, Phys. Rev. Lett. **62**, 981 (1989).

[28] V. A. Ivanov, P. E. Kornilovich, and V. V. Bobryshev, Physica C **235–240**, 2369 (1994).

[29] M. Yu. Kagan and T. M. Rice, J. Phys.: Condens. Matter **6**, 3771 (1994).

[30] R. Micnas, J. Ranninger, and S. Robaszkiewicz, J. Phys. C **21**, L145 (1988).

[31] F. J. Ohkawa and H. Fukuyama, J. Phys. Soc. Jpn. **53**, 4344 (1984).

[32] C. Lacroix, C. Bastide, and A. da Rosa Simoes, Physica B **163**, 124 (1990).

[33] *Superconductivity*, ed. by R. D. Parks, Plenum Press, New York–London (1981).

[34] M. Frick and T. Schneider, Z. Phys. B **78**, 159 (1990).

[35] M. Dressel, O. Klein, G. Grüner *et al.*, Phys. Rev. B **50**, 13603 (1994).

[36] M. Lang, N. Toyota, T. Sasaki, and H. Sato, Phys. Rev. Lett. **69**, 1443 (1992).

[37] Y. Nakazawa and K. Kanoda, Phys. Rev. B **55**, 8670 (1997); Physica C **282–287**, 1897 (1997).

[38] B. T. Geilikman, Usp. Fiz. Nauk **88**, 327 (1966).

[39] G. E. Volovik and L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **88**, 1412 (1985) [Sov. Phys. JETP **61**, 843 (1985)].

[40] Y. Hasegawa, J. Phys. Soc. Jpn. **65**, 3131 (1996).

[41] B. Andraka, G. R. Stewart, K. D. Carbon *et al.*, Phys. Rev. B **42**, 9963 (1990).

[42] Y. Nakazawa, A. Kawamoto, and K. Kanoda, Phys. Rev. B **52**, 12890 (1995).

[43] T. Ichinomiya and K. Yamada, J. Phys. Soc. Jpn. **65**, 1764 (1996).

[44] S. M. de Soto, C. P. Slichter, A. M. Kini *et al.*, Phys. Rev. B **52**, 10364 (1995).

[45] H. Mayaffre, P. Wzietek, D. Jerome *et al.*, Phys. Rev. Lett. **75**, 4122 (1995).

[46] K. Kanoda, K. Miyagawa, A. Kawamoto, and Y. Nakazawa, Phys. Rev. B **54**, 76 (1996).

[47] Yu. V. Kopaev and L. R. Tagirov, JETP Lett. **49**, 499 (1989).

[48] K. A. Moler, D. J. Baar, J. C. Urbach *et al.*, Phys. Rev. Lett. **73**, 2744 (1994).

[49] T. Imai, T. Shimizu, H. Yasuoka *et al.*, J. Phys. Soc. Jpn. **57**, 2280 (1988).

[50] Y. Kitaoka, S. Hiramatsu, Y. Konori *et al.*, Physica C **153**, 83 (1988).

[51] N. Bulut and D. J. Scalapino, Phys. Rev. Lett. **67**, 2898 (1991).

[52] J. P. Lu, Modern Phys. Lett. B **6**, 547 (1992).

[53] J. Caulfield, W. Lubczynski, F. L. Pratt *et al.*, J. Phys.: Condens. Matter **6**, 2911 (1994).

[54] F. L. Pratt, J. Singleton, M. Kurmoto *et al.*, in *The Physics and Chemistry of Organic Superconductors*, ed. by G. Saito and S. Kagoshima, Springer-Verlag, Berlin–Heidelberg (1990), p. 200.

[55] A. Ugawa, G. Ojima, K. Yakushi, and H. Kuroda, Phys. Rev. B **38**, 5122 (1988).

[56] M. V. Kartsovnik, G. Yu. Logvenov, H. Ito *et al.*, Phys. Rev. B **52**, 15715 (1995).

[57] T. Sasaki, H. Sato, and N. Toyota, Physica C **185–189**, 2687 (1991).

[58] C.-P. Heidman, H. Müller, K. Biberacher *et al.*, Synth. Met. **41–43**, 2029 (1991).

[59] J. Wosnitza, G. W. Crabtree, H. H. Wang, *et al.*, Phys. Rev. B **45**, 3018 (1992).

[60] J. Caulfield, J. Singleton, F. L. Pratt *et al.*, Synth. Met. **61**, 63 (1993).

[61] T. Komatsu, N. Matsukawa, T. Inoue, and G. Saito, J. Phys. Soc. Jpn. **65**, 1340 (1993).

Translation provided by the Russian Editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Long-lived excited states and photoluminescence excitation spectra in single crystals of fullerene C_{60}

V. V. Kveder,*) V. D. Negrii, É. A. Shteĭnman, A. N. Izotov, Yu. A. Osip'yan, and R. K. Nikolaev

Solid-State Physics Institute, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 4 August 1997)
Zh. Éksp. Teor. Fiz. **113**, 734–746 (February 1998)

Photoluminescence, optical absorption spectra, and photoluminescence excitation spectra were measured on large (2–3 mm), very pure crystals of fullerene C_{60} at 5 K. It is shown that the main contribution to the photoluminescence of these crystals is from singlet and triplet excitons captured on crystal defects. The concentration of these defects does not exceed 10^{18} cm^{-3}, and the lifetime of triplet excitons on these defects is about 3 ms. It is shown that the symmetry distortion of the C_{60} molecules at the defects is rather large and causes the oscillator strength of the zero-phonon optical transitions to be comparable to the most intense optical transitions with the participation of intramolecular vibrations. © *1998 American Institute of Physics.* [S1063-7761(98)02702-4]

1. INTRODUCTION

The photoluminescence of an isolated C_{60} molecule is determined by optical transitions between the lowest unoccupied molecular orbitals (LUMO) and the highest occupied molecular orbitals (HOMO). According to theoretical calculations (see, for instance, Refs. 1–3), there are three LUMO levels of the singlet electronic excitations of the C_{60} molecule, whose symmetry corresponds to the irreducible representations T_{1g}, T_{2g}, and G_g of the point symmetry group I_h. All these lie about 2 eV above the A_g ground (HOMO) state. The calculated energy positions of the T_{1g}, T_{2g}, and G_g LUMO levels are so close that theory does not make it possible to unambiguously determine in what order they occur. For example, calculations using the ARGUS program (see Ref. 4) give 1.86, 1.90, and 1.93 eV for these energy levels, the computations of Ref. 5 give 2.10, 2.17, and 2.23 eV, and the CNDO/S method was used in Ref. 1 to obtain the values 2.33, 2.29, and 2.34 eV. Besides singlet excited states, there are the levels of the triplet electronic excited states, the lowest of which are, respectively, ${}^3T_{1g}$ and ${}^3T_{2g}$. The triplet levels lie somewhat lower than the singlet levels (by about 0.2 eV). Because the A_g ground state and the lowest excited states T_{1g}, T_{2g}, and G_g have the same parity, single-photon transitions between them are dipole-forbidden. Therefore, optical transitions between the HOMO and LUMO levels in an isolated molecule are possible only when odd-symmetry intramolecular vibrational modes of the C_{60} molecule are excited,[1,6,7] so that ${}^1A_g - {}^1T_{1g}$ transitions are allowed with the participation of a_0-, t_{1u}-, and h_u-vibrations, while ${}^1A_g - {}^1T_{2g}$ transitions are allowed with the participation of g_u- and h_u-vibrations. Thus, the photoluminescence spectra for an isolated molecule are a series of lines with energies $E_{0-0} - \hbar\omega_{\rm vibr}$, while the absorption spectra (and the photoluminescence excitation spectra) are a series of lines with energies $E_{0-0} + \hbar\omega_{\rm vibr}$, where $\omega_{\rm vibr}$ are the frequencies of the corresponding intramolecular vibrational modes and E_{0-0} is the energy difference between the HOMO and the LUMO levels (see, for example, Ref. 8).

The electronic spectrum of crystalline C_{60} must differ from that of an isolated molecule because of intermolecular interactions. In the standard one-electron approximation, the totality of the occupied valence band must be formed from the ${}^1A_g$ ground HOMO level, while the singlet and triplet exciton bands and the conduction band must be formed from the unoccupied LUMO levels. The bottom of the conduction band must of course lie above the bottom of the singlet-exciton band by an amount equal to the exciton binding energy, which can be roughly estimated as 0.3–0.4 eV. The overlap of the electronic wave functions of adjacent molecules in the crystal is small, and the allowed energy bands are rather narrow (about 0.5 eV). This means that electron–electron correlations and polaron effects can have a very substantial effect on the electronic spectrum and the electronic properties of C_{60} crystals. Thus, it is not yet completely clear to what degree crystalline C_{60} can be regarded as an ordinary semiconductor with one-electron energy bands. It is also unclear to what extent the symmetry distortion of the molecules by intermolecular interactions makes the zero-phonon optical 0–0 transitions allowed in an actual crystal.

A rich set of lines is experimentally observed in the photoluminescence spectra of C_{60} crystals in the energy interval from 1.9 to 1 eV. References 4 and 7 ascribed the photoluminescence lines to recombination of free singlet Frenkel excitons with the participation of various intramolecular vibrational modes (see Table I). Reference 9 found that the intensity of a series of lines in the photoluminescence spectrum of C_{60} crystals varies from sample to sample and even from place to place in one sample. Based on this, part of the photoluminescence lines were ascribed to Frenkel excitons localized on crystal defects.[9,10] It was assumed in this case

TABLE I. Energy positions of certain photoluminescence lines in C_{40} crystals and their interaction according to Refs. 4 and 9.

No.	Energy, eV our data	Energy, eV data of Ref.9	Energy, eV data of Ref.4
1	1.815	$1.815 = X1 - t_{1u}(0.178)$	
2			$1.804 = {}^1T_{2g} - h_u(0.090)$
3		$1.784 = X1 - h_u(0.204)$	
4	1.762	$1.762 = X2 - t_{1u}(0.178)$	$1.763 = {}^1T_{2g} - g_u(0.125)$
5	1.728	$1.726 = X2 - h_u(0.204)$	$1.717 = {}^1T_{1g} - t_{1u}(0.150)$
6			$1.697 = {}^1T_{1g} - t_{1u}(0.178)$
7	1.688	$1.693 = C - t_{1u}(0.178)$	$1.689 = {}^1T_{2g} - h_u(0.204)$
8	1.654	$1.659 = C - h_u(0.204)$	
9		$1.661 = X3 - t_{1u}(0.178)$	$1.664 = {}^1T_{1g} - h_u(0.204)$
10	1.636	$1.638 = X4 - t_{1u}(0.178)$	$1.635 = {}^1T_{1g} - h_u(0.066) - a_g(0.179)$
11	1.623		
12	1.607	$1.603 = X5 - t_{1u}(0.178)$	$1.598 = {}^1T_{1g} - h_u(0.090) - a_g(0.179)$
13	1.590		
14	1.571	$1.570 = X5 - h_u(0.178)$	
15	1.549		$1.541 = {}^1T_{1g} - t_{1u}(0.150) - a_g(0.179)$
16	1.528		
17	1.507		$1.510 = {}^1T_{2g} - h_u(0.204) - a_g(0.179)$
18	1.482		
19	1.463		
20	1.445		
21	**1.472**		
22	**1.427**		
23	**1.402**		$1.414 = {}^1T_{1g} - h_u(0.090) - 2a_g(0.179)$
24	**1.379**		
25	**1.352**		
26	**1.330**		$1.337 = {}^1T_{1g} - t_{1u}(0.178) - 2a_g(0.179)$
27	**1.305**		
O_2	**0.964**		

that, as for free molecules, the probability of 0–0 transitions is negligible, and that the observed photoluminescence lines are shifted in energy from the electronic 0–0 transitions by the energy of the h_u and t_{1u} intramolecular vibrations. On the other hand, it was shown in Ref. 11 that part of the photoluminescence lines can be interpreted in terms of zero-phonon 0–0 transitions corresponding to recombination of singlet excitons on deep traps.

The presence of photoluminescence lines associated with the recombination of triplet excitons can be expected in the energy region below 1.6 eV. Measurements of the optical absorption at triplet excitons and experiments in ESR give a relatively short lifetime of these excitons—from 40 μs to several hundred μs,[12,13] which is evidence that nonradiative processes have high efficiency. In this connection, the possibility of observing the luminescence of triplet excitons seems extremely problematic. However, distortion of the symmetry of the molecules (as in the case of singlet excitons) can strongly increase the probability of radiative recombination of triplet excitons. Indeed, in Ref. 14 it was found that triplet phosphorescence is very weak in a solution of C_{60} in methyl cyclohexane but becomes much stronger in a solution of C_{60} in $N'N$-diethyl aniline. As far as we know, the first successful observation of triplet-exciton photoluminescence in C_{60} crystals was reported in Ref. 15, where lines with energies 1.501, 1.467, 1.450, 1.439, and 1.406 eV were detected. These lines were ascribed to the recombination of triplet excitons at two types of traps, with the 1.439- and 1.406-eV lines being ascribed to recombination at deep traps formed by pairs (dimers) of C_{60} molecules.

In order to improve the understanding of the electronic properties of C_{60} crystals, the temporal characteristics of the photoluminescence spectra were experimentally studied for this paper, and the excitation spectra of individual photoluminescence lines and the optical absorption spectra were also measured in large high-purity C_{60} crystals.

2. SAMPLES AND EXPERIMENTAL TECHNIQUE

Single crystals of C_{60} were grown at the Solid-State Physics Institute, Russian Academy of Sciences, by a physical vapor-transport method (sublimation) in an evacuated ampule at a temperature of about 500 °C. The starting C_{60} material, after being purified by chromatography and after traces of solvent were removed by multiple resublimation in vacuum, had a purity no worse than 99.95%.

The photoluminescence spectra of the crystals were recorded by means of a cooled FEU-62 photomultiplier or a cooled germanium photoresistor. The overall spectral sensitivity of the apparatus was calibrated by recording the spectra of an incandescent tungsten lamp. The photoluminescence spectra were then normalized by means of the calibration curve. Thus, the spectra shown in this paper correspond to radiation density on an energy scale: $dN_{\mathrm{ph}}(E)/dE$, where $N_{\mathrm{ph}}(E)$ is the number of radiated photons with energy E.

The photoluminescence was excited by a He–Ne laser ($E_{\mathrm{exc}}=1.959$ eV) with a power of 3 mW, which was attenuated to the necessary power by a rotatable polarizer. To obtain the photoluminescence excitation spectra, it was recorded at a fixed wavelength while excitation was provided by radiation from a halogen lamp transmitted through a fast monochromator. The photoluminescence excitation spectra were normalized to the spectral density of the exciting light, calibrated by means of a thermistor power meter.

To measure the relaxation time of the photoluminescence, the exciting radiation of the laser was modulated at a frequency F by a lithium niobate-based electrooptic modulator, and the photoluminescence spectra were recorded at this frequency by means of a lock-in amplifier. The relaxation time corresponding to specific lines in the photoluminescence spectrum was computed from the dependence of the intensity of these lines on the modulation frequency. This method of determining the excited-state lifetime differs from the conventional method based on directly observing the kinetics of the photoluminescence decrease after a laser pulse and is far more sensitive.

The photoluminescence was measured at a temperature of 5 K. To avoid unwanted effects associated with photostimulated polymerization (dimerization) and other irreversible photostimulated processes in the test samples, all the measurements were made with a laser excitation density no greater than 2.5 mW/mm^2. The photoluminescence spectra of the samples do not change during the measurements in this case and are very reproducible from measurement to measurement.

3. EXPERIMENTAL RESULTS

Figure 1 shows the photoluminescence spectra of one of the samples, recorded with a pump density of 0.25 mW/mm^2 at two modulation frequencies $F=17$ Hz and 306 Hz. It can be seen that the photoluminescence spectrum contains a large number of lines, whose half-width ranges from 7 to 12 meV. The energy positions of the best resolved lines are given in Table I. It should be pointed out that the position and even the width of the photoluminescence lines are reproduced with good accuracy in different samples, but the relative intensity of the different photoluminescence lines can vary appreciably from one batch of crystals to another. The amplitudes of the highest-energy photoluminescence lines (1.815, 1.762, 1.728, and 1.688 eV) vary especially strongly.

The observed photoluminescence lines can be separated into two categories: lines whose amplitudes are independent of F in the frequency range that we used (from 10 Hz to 10 kHz) and lines whose intensities sharply decrease even at modulation frequencies of about 300 Hz. The first group of lines corresponds to short-lived excited states (with lifetimes τ known to be shorter than 10^{-5} sec), whereas the second group corresponds to states with a surprisingly long lifetime (more than 10^{-3} sec). The lines corresponding to long-lived states are distinguished in Table I by boldface.

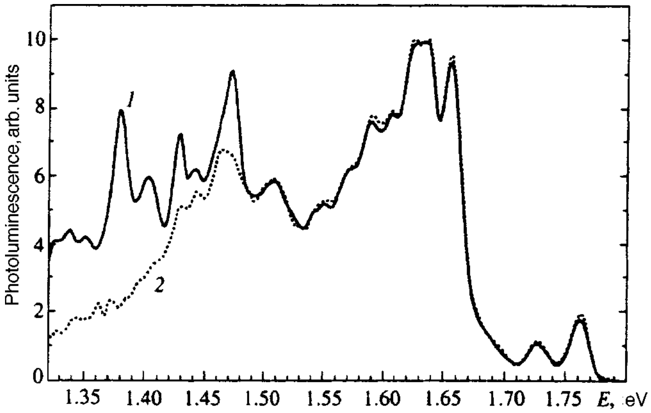


FIG. 1. Photoluminescence spectra of crystal R15, recorded at 5 K and a power of $P=0.25$ mW/mm^2 from a He–Ne laser (632.8 nm) at two modulation frequencies: 17 Hz (continuous curve *1*) and 306 Hz (dotted curve *2*).

The photoluminescence spectra recorded at modulation frequencies $F>500$ Hz correspond only to short-lived states, while the photoluminescence spectra of the long-lived states can be obtained by subtracting the photoluminescence spectra recorded at low ($F<20$ Hz) and high ($F>500$ Hz) modulation frequencies. Figure 2 also shows the difference spectra measured at several laser-excitation densities. The intensity of the spectra is normalized to the excitation density. Unlike the photoluminescence lines of the short-lived states, which virtually do not saturate in the range of excitation power that we investigated (from 0.05 to 3 mW/mm^2), the photoluminescence lines of the long-lived states quickly saturate, and their relative contribution to the photoluminescence spectrum is small at high excitation powers. Figure 3 shows the experimental dependences of the intensity of cer-

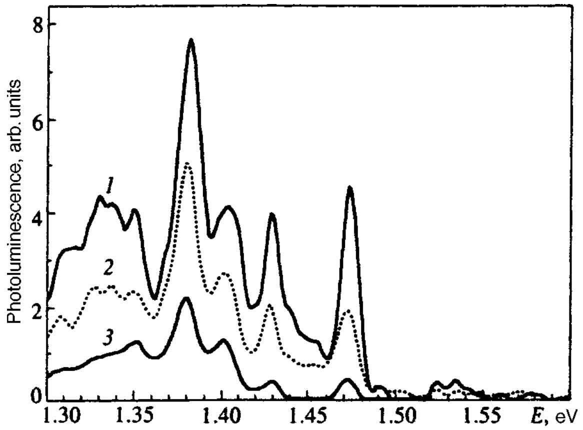


FIG. 2. Photoluminescence spectra of the long-lived states of crystal R15, obtained by subtracting the spectra recorded at $F=17$ Hz and 570 Hz; $T=5$ K, excitation by a He–Ne laser, spectrum *1* corresponds to $P=0.1$ mW/mm^2, *2*—$P=0.4$ mW/mm^2, *3*—$P=2.3$ mW/mm^2.

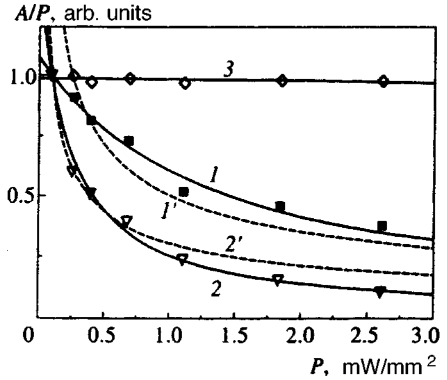


FIG. 3. Total intensity A of the photoluminescence line divided by excitation power P vs excitation power from a He–Ne laser at $T=5$ K; *1*—long-lived lines 1.402 and 1.379 eV, *2*—long-lived lines 1.472 and 1.427 eV, *3*—short-lived lines. The points show the experimental data, the continuous curves are calculated in the tau approximation, and the dashed curves (1′ and 2′) are calculated for a two-particle reaction.

tain long-lived photoluminescence lines on excitation power P. The intensities are normalized to the excitation power and are reduced to unity at $P=0.1$ mW/mm^2 to make them easy to compare. For comparison, Fig. 3 also presents data for the short-lived lines (points *3*), which show that there is no saturation within the limits of accuracy of the measurement.

The continuous curves in Fig. 3 are calculated in the simplest model, which assumes that the excited-state lifetime τ is independent of the excitation intensity (the tau approximation). In this case, the number N_{PL} of excited molecules that contribute to a given photoluminescence line is described by the simple kinetic equation

$$\partial N_{PL}/\partial t=G(1-N_{PL}/N_0)-N_{PL}/\tau, \tag{1}$$

where N_0 is the total number of C_{60} molecules that can contribute to the given photoluminescence line, τ is the excited-state lifetime, and G is the generation rate, which is proportional to the power P of the laser radiation, $G=aP$. Equation (1) has a simple solution:

$$N_{PL}(t)=G/(G/N_0+1/\tau)+(N_{PL(t=0)}-G/(G/N_0+1/\tau_0))\times\exp(-(G/N_0+1/\tau)t). \tag{2}$$

The photoluminescence intensity is proportional to $N_{PL}\tau/\tau_R$, where τ_R is the radiative recombination time of the excited state. In order to compute the experimentally measured signal A at the output of the lock-in amplifier at a given modulation frequency F, Eq. (2) was substituted into an integral equation corresponding to the transfer function of the lock-in amplifier:

$$A\propto\int_0^{1/2F}N_{PL}(t,G=G_0)dt-\int_{1/2F}^{1/F}N_{PL}(t,G=0)dt, \tag{3}$$

and the steady-state amplitude A for repetition of the modulation cycles was computed self-consistently. In this case, the adjustable parameters that determine the calculated dependences of A on P and on F were the quantities τ and a/N_0.

As can be seen from Fig. 3, the theoretical curves calculated in the tau approximation agree well with the experimental data. In this model, the photoluminescence signal saturates because the fill factor of the excited states $f=N_{PL}/N_0$ approaches unity when $G>N_0/\tau$. This is possible only if one is dealing with the photoluminescence of excitons trapped at defects of the crystal lattice, i.e., when the number N_0 is much less than the total number of molecules in the crystal. In the case of free excitons, N_0 almost corresponds to the total number of molecules in the crystal, and $N_{PL}\ll N_0$. In this case, saturation of the photoluminescence could be explained on the assumption that, as N_{PL} increases, the probability of nonradiative recombination increases because of some two-particle processes:

$$\partial N_{PL}/\partial t=G-N_{PL}/\tau_0-\alpha N_{PL}^2. \tag{4}$$

Theoretical curves calculated from Eq. (4) are shown in Fig. 3 by dashed curves. It can be seen that they describe the experimental behavior significantly worse than the tau approximation.

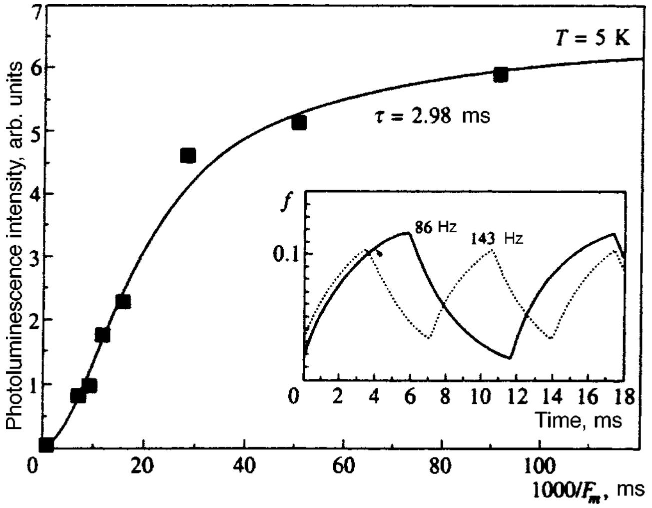


FIG. 4. Total intensity of long-lived photoluminescence lines with energies 1.402 and 1.379 eV vs modulation frequency F at $T=5$ K. The points show experimental data, and the continuous curve is calculated for $\tau=2.98$ ms. The inset shows the calculated time dependence of the photoluminescence intensity at two modulation frequencies.

Figure 4 shows the dependence of the photoluminescence intensity of the long-lived states on the modulation frequency F. The theoretical curve calculated in the tau approximation gives a good description of the experimental dependence for values of the adjustable parameter τ around 3 ms.

To obtain additional information on the nature of the photoluminescence, we measured the excitation spectra for individual photoluminescence lines. The results are shown in Fig. 5. Curve *3* corresponds to the photoluminescence spectra recorded with excitation by a He–Ne laser ($E=1.959$ eV), while curves *1*, *2*, *4*, and *5* correspond to the excitation spectra of the photoluminescence lines with energies of 1.688, 1.654, 1.626, and 1.379 eV, respectively. To make it easy to compare them, the excitation spectra are reduced to unity at an excitation energy of $E=1.959$ eV, corresponding to the He–Ne laser.

Figure 6 shows the absorption spectrum of the same crystal (curve *A*). For comparison, Fig. 6 also shows the

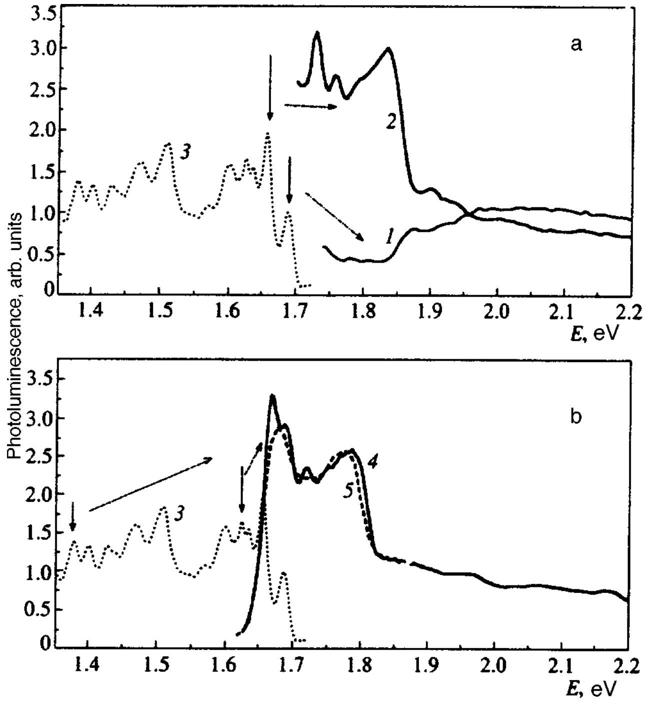


FIG. 5. Photoluminescence excitation spectra, measured for various photoluminescence lines in crystal R5 at $T=10$ K. Dotted curve *3* shows the photoluminescence spectrum with excitation by a He–Ne laser ($E=1.959$ eV), and curves *1, 2, 4* and *5* show the excitation spectra for the photoluminescence lines with energies 1.688, 1.654, 1.626, and 1.379 eV, respectively (see arrows on the photoluminescence spectra). The excitation spectra are reduced to unity at $E=1.959$ eV.

photoluminescence excitation spectrum of the lines with energies 1.688 and 1.626 eV. As can be seen, a rather strong absorption edge (above $200\ \mathrm{cm}^{-1}$) corresponds to $E=1.82$ eV. This energy almost coincides with the step in the excitation spectrum of the 1.688-eV photoluminescence line. However, some absorption (at a level of about $10-30\ \mathrm{cm}^{-1}$) is also observed below this edge, and this has a pronounced structure in the form of overlapping lines whose position coincides with the position of the lines in the photoluminescence excitation spectrum. It should be pointed out that the indicated features are observed in all the crystals that were studied, but that the value of the absorption in the region $E<1.82$ eV varies from sample to sample.

Figure 7 shows the absorption spectra of a C_{60} crystal at various temperatures. The variations of the absorption spectra are insignificant in the temperature interval from 5 to 80 K. An appreciable decrease of the energy is observed at temperatures above 90 K, corresponding to the strong absorption edge, whereas the weak absorption lines (in the 1.65–1.76-eV region) almost do not shift but only broaden slightly.

4. DISCUSSION OF THE RESULTS

There is currently no doubt that the short-lived photoluminescence of C_{60} crystals is caused by singlet Frenkel excitons. Many investigators have assumed that the selection rules for optical transitions are the same in crystalline C_{60} as in an isolated molecule, i.e., that optical transitions between the electronic HOMO and LUMO states are possible only with excitation of the corresponding intramolecular vibrations. Table I shows the interpretation of certain photoluminescence lines in C_{60} crystals according to the data of Refs. 4 and 9. It is assumed in those papers that the energy E_{0-0} of the free singlet ${}^1T_{1g}$ exciton equals 1.871 eV, while all the observable photoluminescence lines correspond to transitions with the participation of vibrational molecular modes. In Ref. 4, all the photoluminescence lines are ascribed to free excitons, whereas, in Ref. 9, some of the lines are ascribed to excitons bound at shallow exciton traps. Following this interpretation, it can be expected that, even if polaron effects are neglected, the photoluminescence excitation spectrum of free excitons must start with an energy of about 1.871 eV $+\hbar\omega_{\mathrm{vibr}}$ [i.e., with an energy of about 1.96 eV for one of the softest active modes h_u (0.09 eV)].

As can be seen from Fig. 5, for the 1.688-eV photoluminescence line, there actually is a substantial difference (of about 150 meV) between the step at $E=1.84$ eV in the photoluminescence excitation spectrum and the position of the line. In this case, the step of the excitation spectrum of this line virtually coincides with the edge of the ''strong'' optical absorption of the crystals (see Fig. 6).

We assume that the strong absorption edge corresponds to the excitation of free singlet excitons in the crystal. The difference of 150 meV between the 1.688-eV photoluminescence line and the step in its excitation spectrum could be

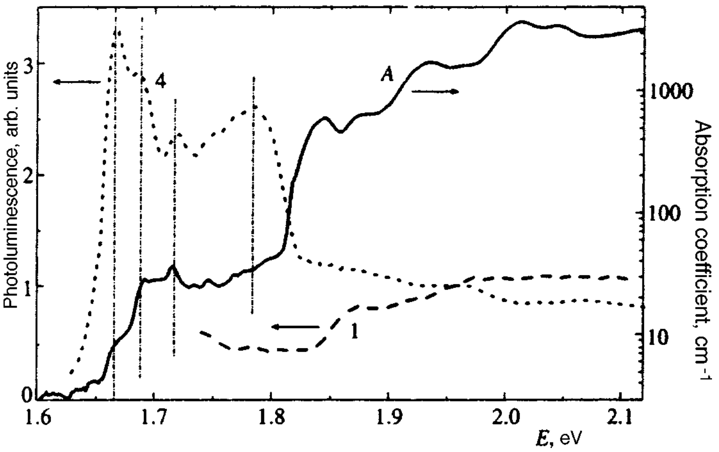


FIG. 6. Absorption spectrum of crystal R5 at $T=10$ K (continuous curve *A*). Curves *1* and *4* show the photoluminescence excitation spectra of the lines with energies 1.688 and 1.626 eV, reduced for comparison.

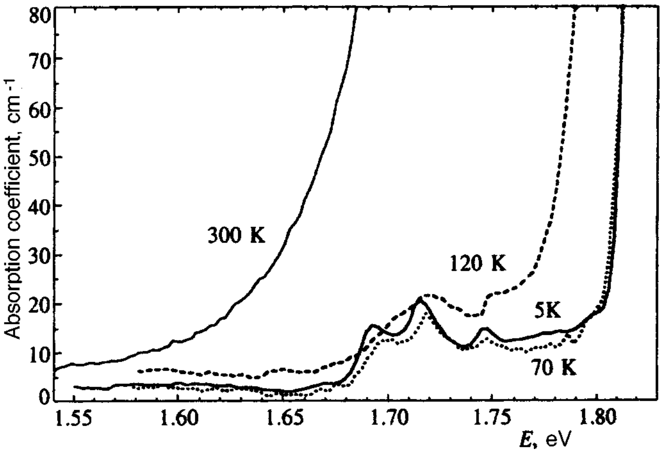


FIG. 7. Absorption spectrum of crystal R14, measured at $T=5$, 70, 120, and 300 K.

ascribed to excitation of the corresponding vibronic modes if it is assumed that the line with energy 1.688 eV also corresponds to free singlet excitons. This would mean that zero-phonon transitions are, as before, forbidden for an ''ideal'' (defect-free) crystal.

However, it is possible that the line with energy 1.688 eV corresponds not to free excitons but to excitons at shallow traps. Then it is possible that the indicated difference in energies (150 meV) can correspond to the binding energy of an exciton with these traps and to the polaron effect (i.e., an energy decrease due to deformation of the molecule or the crystal lattice by an exciton). Thus, we cannot exclude that both the line with energy 1.688 eV and the strong absorption edge correspond to zero-phonon optical transitions (it is possible that ordinary crystal phonons, which are very shallow in C_{60} molecules, participate here). In this case, the strong absorption lines with energies 1.841 and 1.876 eV (see Fig. 6) can be ascribed to zero-phonon optical transitions from the ${}^1A_{1g}$ ground state to the $k\approx 0$ states of the singlet-exciton bands formed from the ${}^1T_{1g}$ and ${}^1T_{2g}$ levels (the difference of about 35 meV between them agrees with the calculated values[1,4,5]). Then the absorption lines with energies 1.932 and 1.966 eV can correspond to a vibronic sideband of the first two lines with the participation of the h_u mode (0.09 eV), while the 2.010 and 2.047-eV lines correspond to a vibronic sideband with participation of other modes with an energy of about 170 meV [for example, t_{1u} (0.178 eV) or h_u (0.157 eV)].

The excitation spectra of the photoluminescence lines beginning with an energy of 1.65 eV and below are very similar to each other and sharply differ from the excitation spectra of the photoluminescence lines with higher energy (see Fig. 5). The experimental data make it possible to assume that these lines correspond to excitons at deep exciton traps caused by crystal-lattice defects. In this case, all the short-lived states evidently correspond to singlet excitons. The very small difference between the edge of the excitation spectrum and the position of the photoluminescence line, 1.65 eV and 1.63 eV, means that at least these lines correspond to zero-phonon transitions, i.e., that the deformation of C_{60} molecules in these defects because of intermolecular interactions is rather small. Then the group of short-lived photoluminescence lines in the energy region below 1.55 eV can be regarded as a vibronic sideband of the lines in the region with energy 1.7–1.6 eV involving a vibrational mode with an energy of about 0.18 eV. These conclusions correlate with those of Ref. 16, where is was also shown that deep exciton traps can make a large contribution to the photoluminescence of C_{60} crystals and that the probability of zero-phonon transitions can be large for these traps.

Following Ref. 15, we assume that the long-lived excitations that we observed in the photoluminescence correspond to triplet excitons localized on crystal-lattice defects. The photoluminescence spectra of triplet excitons in our crystals differ from those observed in Ref. 15. The lifetimes of these excitations also strongly differ from those in Ref. 15: in our case, $\tau\approx 3$ ms, whereas $\tau\approx 0.4$ ms in Ref. 15. This means that the defects that dominate our samples differ somewhat from those observed in Ref. 15. Starting from the data on the saturation of the long-lived photoluminescence, an upper limit for the concentration of these defects in our samples can be estimated. In the tau approximation, saturation of the photoluminescence corresponds to the situation in which the concentration of triplet excitons captured on those defects that possess a large ratio τ/τ_R becomes comparable with the concentration of these defects. Recalling that this occurs at a laser power of about 1 mW/mm^2, while the penetration depth of the He–Ne-laser radiation is about 10 μm, we find that the defect concentration does not exceed 3×10^{18} cm^{-3} even if it is assumed that all the excitons generated by the laser radiation are captured only on the indicated defects.

Thus, the concentration of deep traps discussed here does not exceed 10^{-3}–10^{-4} of the total number of C_{60} molecules. In this connection, one needs to explain why, with such a small concentration of defects, they virtually completely determine the photoluminescence spectrum of the test crystals. Indeed, the photoluminescence from the deep exciton traps is dominant in our crystals not only with excitation by a He–Ne laser ($E=1.96$ eV), but also with excitation by an argon laser having an energy of $E=2.41$ eV. As follows from the absorption spectra, the fraction of photons absorbed directly by defects is very small in both cases (especially in the case of the argon laser). Therefore, the only reasonable explanation of this fact is to assume that free excitons in C_{60} crystals are very mobile and possess a rather large diffusion coefficient even at a temperature of 5 K. In this case, the overwhelming majority of free excitons manage to be captured on deep traps during their lifetime, and excitons at defects therefore dominate in the photoluminescence spectrum.

If it is not assumed that the excitons have high mobility, it has to assumed that the luminescence quantum yield at defects is greater by three orders of magnitude than the quantum yield from C_{60} molecules found in an ideal crystalline neighborhood. Of course, the quantum yield can be expected to be enhanced at defects because of the removal of the forbiddenness on zero-phonon transitions, but it is hard to suppose that the quantum yield can increase by three orders of magnitude in this case.

Recalling that the crystals studied for this paper were fabricated from very pure raw material and were subjected to multiple resublimation in vacuum, we assume that the defects under discussion are intrinsic defects of the crystal and are not associated with impurities. It is possible that these defects result from plastic deformation of the crystals during growth under the action of their own weight, since C_{60} crystals are very plastic at the growth temperature (about 450–470 °C).

5. CONCLUSION

This paper has discussed comparatively large (2–3 mm across) pure C_{60} crystals grown by resublimation in vacuum. It has shown that the photoluminescence spectra in such C_{60} crystals mainly corresponds to excitons captured at deep traps caused by crystal-lattice defects. The atomic concentration of defects is at most 10^{-3}–10^{-4}, and the fact that they dominate in the photoluminescence means that the free excitons have high mobility even at 5 K. The very large contribution to the photoluminescence at low excitation density comes from long-lived excitations with a lifetime of about 3 ms, which probably corresponds to triplet excitons captured at crystal defects. A comparison of the photoluminescence spectra and their excitation spectra shows that, at least in the case of crystal defects, distortion of the symmetry of C_{60} molecules due to intermolecular interactions removes the forbiddenness on zero-phonon optical transitions.

This work was carried out under the ''Fullerenes and atomic clusters'' project of the federal program ''Critical directions in the physics of condensed media,'' financed by the Ministry of Science of Russia.

*⁾e-mail: kveder@issp.ac.ru

---

[1] F. Negri, G. Orlandi, and F. Zerbetto, J. Chem. Phys. **97**, 6496 (1992).
[2] I. Laszlo and L. Udvardi, Chem. Phys. Lett. **136**, 418 (1987).
[3] K. Yabana and G. F. Bertsch, Chem. Phys. Lett. **197**, 32 (1992).
[4] M. Diehl, J. Degen, and H.-H. Schmidtke, J. Chem. Phys. **99**, 10092 (1995).
[5] R. D. Bendale, J. F. Stanton, and M. C. Zerner, Chem. Phys. Lett. **194**, 467 (1992).
[6] F. Negri, G. Orlandi, and F. Zerbetto, Chem. Phys. Lett. **144**, 31 (1988).
[7] Y. Wang *et al.*, Phys. Rev. B **51**, 4547 (1995).
[8] A. Sassara, G. Zerra, and M. Chergui, J. Phys. B **29**, 4997 (1996).
[9] W. Guss, J. Feldmann, E. O. Goebel, C. Taliani, H. Mohn, W. Muller, P. Haussler, and H.-U. ter Meer, Phys. Rev. Lett. **72**, 2644 (1994).
[10] V. D. Negri, V. V. Kveder, Yu. A. Ossipyan, I. N. Kremenskaya, and R. K. Nikolaev, Phys. Status Solidi B **199**, 587 (1997).
[11] D. J. van den Heuvel *et al.*, J. Phys. Chem. **99**, 11644 (1995).
[12] M. R. Wasilewski, M. P. O'Neil, K. R. Lykke, M. J. Pellin, and D. M. Gruen, J. Am. Chem. Soc. **113**, 2774 (1991).
[13] J. M. Arbogast *et al.*, J. Phys. Chem. **95**, 11 (1991).
[14] Y. Wang, J. Phys. Chem. **96**, 764 (1992).
[15] D. J. van den Heuvel, I. Y. Chan, E. J. J. Groenen, J. Schmidt, and G. Meijer, Chem. Phys. Lett. **231**, 111 (1994).
[16] D. J. van den Heuvel, I. Y. Chan, E. J. J. Groenen, M. Matsushita, J. Schmidt, and G. Meijer, Chem. Phys. Lett. **233**, 284 (1995).

Translated by W. J. Manthey

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 86, NUMBER 2 FEBRUARY 1998

Self-consistent solutions of the problem of the interaction of a two-frequency field with a system of three-level atoms in the form of a phase-modulated simulton and a Raman soliton

A. V. Andreev*)

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

(Submitted 28 April 1997)
Zh. Éksp. Teor. Fiz. **113**, 747–762 (February 1998)

This paper discusses features of the interaction of a system of three-level atoms with a bichromatic electromagnetic field. Steady-state solutions are found in the form of unimodal and nonunimodal solitary waves corresponding to a pair of pulses in the form of a simulton or a Raman soliton. An analytic form of the solutions is found, corresponding to a phase-modulated simulton and a Raman soliton. The conditions are determined for exciting them. The resulting solutions are of interest from the standpoint of the mechanism for forming the two-frequency coupled excitations observed in various experiments. © *1998 American Institute of Physics.* [S1063-7761(98)02802-9]

1. INTRODUCTION

The study of nonsteady-state coherent processes in a system of three-level atoms has continued to interest investigators for many years. This interest has both purely theoretical and applied aspects. The applied aspect is associated with the successful development of systems for converting the frequency and shape of pulses, i.e, for obtaining pulses with high contrast and a frequency shift upward or downward. The discovery of simulton solutions served as a spur to extend the methods of the inverse scattering problem to three-level media.[1] Significant interest in the study of three-level media was associated with the development of inversionless amplification systems[2] and with the experimental and theoretical study of the coherent population-capture effect.[3] Successful experiments on the observation of Raman solitons, carried out recently in various laboratories,[4–7] have caused increased interest in developing a consistent theory of the indicated phenomena.[5,8,9]

Most of the models used in theoretically analyzing the dynamics of the formation of Raman solitons are based on a treatment of two-frequency interaction in a medium of two-level atoms with a dipole-forbidden transition. These models are probably adequate to explain the experiments on stimulated Raman scattering in hydrogen,[4–6] in which the frequency of the pump pulse is far from resonance with any transition. However, in experiments with NH_3 gas,[7] the radiation frequency of the CO_2 laser used as a pump pulse is in quasi-resonance with the vibrational frequency, while the frequency of the Stokes quantum is in quasi-resonance with the rotational transitions of the NH_3 molecule. In this case, the parameters of the soliton strongly depend on the value of the frequency offset, and a consistent model must take into account the population dynamics both at the pump transition and at the Stokes transition. It should be pointed out that simulton solutions were found by the methods of the inverse scattering problem only when the adjacent transitions were exactly in resonance and had equal oscillator strengths.[1,10] In this paper, we find solutions in the form of steady-state pulses when the oscillator strengths at adjacent transitions are different, and the pump-pulse frequency is arbitrarily offset from the frequency of the corresponding transition.

2. BASIC EQUATIONS

The interaction Hamiltonian of an extended system of atoms with a polychromatic field has the form

$$H=H_a+H_{\mathrm{int}}+H_f, \tag{1}$$

where H_a is the Hamiltonian of a free system of atoms,

$$H_a=\sum_{i=1}^{N}\sum_k \hbar\omega_k b_{ki}^+ b_{ki},$$

H_{int} is the interaction Hamiltonian,

$$H_{\mathrm{int}}=-\frac{1}{c}\int \mathbf{j}\cdot\mathbf{A}dV,$$

and H_f is the Hamiltonian of the free field. The current-density operator $\mathbf{j}$ has the form

$$\mathbf{j}=\sum_{i=1}^{N}\sum_k\sum_l \mathbf{m}_{kl}b_{ki}^+ b_{li}\delta(\mathbf{r}-\mathbf{r}_i),$$

where $\mathbf{m}_{kl}=i\omega_{kl}\mathbf{d}_{kl}$ is the matrix element of the transition current between the kth and lth levels of the atom, and $\mathbf{d}_{kl}$ is the dipole moment of the transition.

Let us consider the interaction of the three-level atom whose energy-level diagram is shown in Fig. 1 with a bichromatic field. Let transition $1\leftrightarrow 2$ be dipole-forbidden, $\mathbf{d}_{12}=0$. This case corresponds to the so-called Λ system. However, as is well known, these results are easily generalized to the case of the V system and the Ξ system, which will be done below.

The vector potential $\mathbf{A}(\mathbf{r},t)$ of the electromagnetic field can conveniently be represented in the form

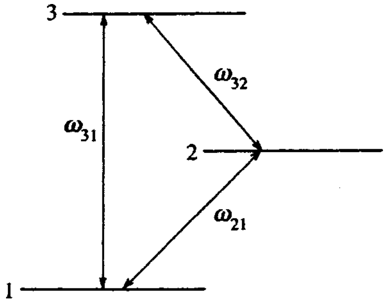


FIG. 1. Energy-level diagram.

$$\mathbf{A}(z,t)=\mathbf{A}_1^+(z,t)\exp\left[i\omega_{31}\left(1-\frac{z}{c}\right)\right]$$
$$+\mathbf{A}_2^+(z,t)\exp\left[i\omega_{32}\left(1-\frac{z}{c}\right)\right]+\text{c.c.} \qquad (2)$$

Using the Hamiltonian of Eq. (1) and assuming that $\mathbf{A}_{1,2}^+(z,t)$ are slowly varying amplitudes, it is not difficult to obtain the following system of equations:

$$\frac{\partial a_1}{\partial t}+\frac{\partial a_1}{\partial z}=\beta_1 b_3^+ b_1, \quad \frac{\partial b_{1,2}}{\partial t}=b_3 a_{1,2},$$
$$\frac{\partial a_2}{\partial t}+\frac{\partial a_2}{\partial z}-i\delta a_2=\beta_2 b_3^+ b_2, \quad \frac{\partial b_3}{\partial t}=-b_1^+ a_1-b_2^+ a_2, \qquad (3)$$

where

$$a_i=\frac{i\tau}{\hbar c}\,\mathbf{m}_{3i}^*\cdot\mathbf{A}_i^+,$$

$$b_i=\frac{1}{\sqrt{N/V}}\frac{1}{V_1}\sum_{n\in V_1}\langle\hat{b}_{\text{in}}\rangle\exp\left[i\omega_i\left(t-\frac{z_n}{c}\right)\right],$$

$$\beta_i=\frac{2\pi\omega_{3i}|\mathbf{d}_{3i}|^2}{\hbar}\frac{N}{V}\tau^2, \qquad \delta=2\pi(\chi(\omega_{32})-\chi(\omega_{31}))\omega_{32}\tau.$$

The dimensionless vector potential $a_{1,2}$ in Eq. (3) is normalized in such a way that the energy density of the field has the form

$$w_f=\left(\hbar\omega_{31}\frac{|a_1|^2}{\beta_1}+\hbar\omega_{32}\frac{|a_2|^2}{\beta_2}\right)\frac{N}{V},$$

where N/V is the density of atoms. After introducing the probability amplitude wave of exciting the ith level, we averaged over the physically small volume V_1. We have introduced dimensionless time $t'=t/\tau$ and position $z'=z/L$ into Eqs. (3), where L is the length of the medium and $\tau=L/c$. The primes in Eqs. (3) and the succeeding expressions have been omitted. The parameter δ in Eqs. (3) takes into account the dispersion, i.e., the difference of the polarizabilities $\chi(\omega_{3i})$ at frequencies ω_{3i}. It is easy to see that by using the transformations $a_2'=a_2e^{-i\delta z}$ and $b_2'=b_2e^{-i\delta z}$, we can eliminate the parameter δ from Eqs. (3), and therefore we set $\delta=0$ in what follows.

The system of Eqs. (3) is a system of equations for the slowly varying field amplitudes that satisfy the condition $|\partial a_i/\partial t|\ll\omega_{3i}|a_i|$ and the probability amplitude b_i of the excitation of the ith level. One can go from the quantum Eqs. (3) to the semiclassical equations by using semiclassical uncouplings of the form $\langle b_i^+a_k\rangle=\langle b_i^+\rangle\langle a_k\rangle$ (see Ref. 11 for more detail).

3. INTEGRALS OF THE MOTION

Let us find some of the integrals of the motion of the system of Eqs. (3), which we shall use in obtaining analytic solutions.

We represent the amplitudes of the waves in Eqs. (3) as

$$a_i(z,t)=A_i(z,t)\exp[i\varphi_i(z,t)],$$
$$b_i(z,t)=B_i(z,t)\exp[i\psi_i(z,t)],$$

and then the system of Eqs. (3) assumes the form

$$\dot{A}_1+A_1'=\beta_1B_3B_1\cos(\psi_1-\psi_3-\varphi_1),$$
$$\dot{A}_2+A_2'=\beta_2B_3B_2\cos(\psi_2-\psi_3-\varphi_2),$$
$$\dot{B}_3=-B_1A_1\cos(\psi_1-\psi_3-\varphi_1)-B_2A_2$$
$$\times\cos(\psi_2-\psi_3-\varphi_2),$$
$$\dot{B}_2=B_3A_2\cos(\psi_2-\psi_3-\varphi_2).$$
$$\dot{B}_1=B_3A_1\cos(\psi_1-\psi_3-\varphi_1),$$
$$(\dot{\varphi}_1+\varphi_1')A_1=\beta_1B_3B_1\sin(\psi_1-\psi_3-\varphi_1), \qquad (4)$$
$$(\dot{\varphi}_2+\varphi_2')A_2=\beta_2B_3B_2\sin(\psi_2-\psi_3-\varphi_2),$$
$$\dot{\psi}_3B_3=-B_1A_1\sin(\psi_1-\psi_3-\varphi_1)-B_2A_2$$
$$\times\sin(\psi_2-\psi_3-\varphi_2),$$
$$\dot{\psi}_2B_2=B_3A_2\sin(\psi_2-\psi_3-\varphi_2),$$
$$\dot{\psi}_1B_1=-B_3A_1\sin(\psi_1-\psi_3-\varphi_1).$$

The partial derivative with respect to time in Eqs. (4) is denoted by a dot, and that with respect to the spatial coordinate is denoted by a prime.

The system of Eqs. (4) possesses the following integrals of the motion and conservation laws:

$$B_1^2(z,t)+B_2^2(z,t)+B_3^2(z,t)$$
$$=B_1^2(z,0)+B_2^2(z,0)+B_3^2(z,0), \qquad (5)$$

$$\dot{\psi}_1(z,t)B_1^2(z,t)+\dot{\psi}_2(z,t)B_2^2(z,t)=\dot{\psi}_3(z,t)B_3^2(z,t), \qquad (6)$$

$$(\dot{\varphi}_i(z,t)+\varphi_i'(z,t))A_i^2(z,t)=-\beta_i\dot{\psi}_i(z,t)B_i^2(z,t), \qquad (7)$$

$$\frac{\partial A_i^2(z,t)}{\partial t}+\frac{\partial A_i^2(z,t)}{\partial z}=\beta_i\frac{\partial B_i^2(z,t)}{\partial t}. \qquad (8)$$

Equation (5) is the law of conservation of the number of excitations or of the number of atoms in the system. Equation (6) can be interpreted as the law of conservation of angular momentum. Equations (7) have the same meaning. Equations (8) are continuity equations. Indeed, adding Eqs. (8) for $i=1,2$, we get

$$\frac{\partial}{\partial t}\left(\frac{|a_1|^2}{\beta_1}+\frac{|a_2|^2}{\beta_2}+1-|b_1|^2-|b_2|^2\right)+\frac{\partial}{\partial z}\left(\frac{|a_1|^2}{\beta_1}+\frac{|a_2|^2}{\beta_2}\right)=0.$$

As we pointed out above, $|a_1|^2/\beta_1$ is the number density of quanta of field A_1, normalized to the number density of atoms N/V, while $1-|b_1|^2-|b_2|^2=|b_3|^2$. Consequently, variation of the sum of the number of the quanta of fields A_1 and A_2 and of the number of atoms on level 3 in volume V is associated only with a flux of quanta through the surface that bounds volume V,

$$\frac{\partial}{\partial t}\int_{z_1}^{z_2}\left(\frac{|a_1|^2}{\beta_1}+\frac{|a_2|^2}{\beta_2}+|b_3|^2\right)dz=\left(\frac{|a_1|^2}{\beta_1}+\frac{|a_2|^2}{\beta_2}\right)\Bigg|_{z=z_1}-\left(\frac{|a_1|^2}{\beta_1}+\frac{|a_2|^2}{\beta_2}\right)\Bigg|_{z=z_2}.$$

4. THE RESONANCE CASE

Let us first consider the resonance case, for which

$$\dot{\varphi}_i=\dot{\psi}_i=0. \tag{9}$$

In this case, the system of Eqs. (4) assumes the form

$$\begin{aligned}&\dot{A}_1=A_1'=\beta_1B_3B_1,\quad \dot{B}_1=B_3A_1,\quad \dot{B}_2=B_3A_2,\\&\dot{A}_2+A_2'=\beta_2B_3B_2,\quad \dot{B}_3=-B_1A_1-B_2A_2.\end{aligned} \tag{10}$$

Such solutions of the system of Eqs. (4), in which amplitudes $A_{1,2}$ depend only on the delay time

$$A_i(z,t)=A_i\left(t-\frac{z}{v}\right),$$

correspond to self-consistent propagation of pulses, where v is the velocity of the self-consistent propagation of the pulses.

In problems involving the interaction of a two-frequency field with a system of three-level atoms, two main types of self-consistent solutions are possible. In the first, the fields on both adjacent transitions have the form of solitary pulses whose amplitude goes to zero as $t\to\pm\infty$. Such solutions are called simultons. In the second, as $t\to-\infty$, the amplitude of one of the fields tends to a constant, while the amplitude of the second field tends to zero. The field that has nonzero amplitude as $t\to-\infty$ acts as a pump pulse. Such solutions are called Raman solitons.

For self-consistent pulse propagation, Eqs. (8) assume the form

$$B_i^2(z,t)+\frac{1}{\beta_i}\left(\frac{1}{v}-1\right)A_i^2\left(t-\frac{z}{v}\right)=B_i^2(z,t_0)+\frac{1}{\beta_i}\left(\frac{1}{v}-1\right)A_i^2\left(t_0-\frac{z}{v}\right), \tag{11}$$

where t_0 is an arbitrary instant.

As can be seen from Eqs. (11), various dynamics of the self-consistent pulse propagation are possible, depending on the sign of the coefficient $k_i=(1-v)/\beta_iv$. For the Λ system, the frequencies ω_{3i} are positive, and therefore the coefficients β_i are positive. Consequently, parameter k_i is positive when $v<1$ and negative when $v>1$. For the V system, coefficients $\beta_i<0$, and parameter k_i is positive when $v>1$ and negative when $v<1$. The normalized pulse velocity is less than unity when the pulse propagates in an initially absorbing medium. This means that, as $t\to\pm\infty$, the conditions $|b_3(t\to\pm\infty)|<|b_2(t\to\pm\infty)|<|b_1(t\to\pm\infty)|$ must be satisfied for the Λ system and the reverse inequalities for the V system. In initially excited media, the velocity of a steady-state pulse can be greater than that of light. In what follows, we shall consider initially unexcited media. Consequently, we assume that $k_i>0$ for the Λ system and $k_i<0$ for the V system. However, these results can easily be generalized to initially excited media.

(a) The Λ system

We first consider the case $k_i>0$. In this case, the right-hand side of Eqs. (11) is positive, and we can rewrite them as

$$B_i^2+k_iA_i^2=C_i^2,$$

or

$$A_i=\gamma_i\sqrt{C_i^2-B_i^2}. \tag{12}$$

where

$$\begin{aligned}&C_i^2=B_i^2(z,t_0)+\frac{1}{\beta_i}\left(\frac{1}{v}-1\right)A_i^2\left(t_0-\frac{z}{v}\right),\\&\gamma_i=\sqrt{\frac{1}{k_i}}=\sqrt{\frac{\beta_iv}{1-v}}.\end{aligned} \tag{13}$$

Using Eqs. (12) and (5), it is easy to obtain the following closed system of equations for amplitudes B_1 and B_2:

$$\begin{aligned}&\dot{B}_1=\gamma_1\sqrt{C_1^2-B_1^2}\sqrt{1-B_1^2-B_2^2},\\&\dot{B}_2=\gamma_2\sqrt{C_2^2-B_2^2}\sqrt{1-B_1^2-B_2^2}.\end{aligned} \tag{14}$$

Let us introduce the new variables $x_{1,2}$,

$$\begin{aligned}&B_1=C_1\cos x_1,\quad B_2=C_2\cos x_2,\\&A_1=\gamma_1C_1\sin x_1,\quad A_2=\gamma_2C_2\sin x_2.\end{aligned} \tag{15}$$

The equations of motion for $x_{1,2}$ have the form

$$\begin{aligned}&\dot{x}_1=-\gamma_1\sqrt{1-C_1^2\cos^2x_1-C_2^2\cos^2x_2},\\&\dot{x}_2=-\gamma_2\sqrt{1-C_1^2\cos^2x_1-C_2^2\cos^2x_2}.\end{aligned} \tag{16}$$

It is easy to see that the variables $x_{1,2}$ are connected by the relationship

$$x_1-x_{10}=\frac{\gamma_1}{\gamma_2}(x_2-x_{20}).$$

Consequently, the coupled motion of the self-consistent pulses at transitions $3\leftrightarrow2$ and $3\leftrightarrow1$ is determined by the following equation of motion for the variable $x=x_1$:

$$\dot{x}=-\gamma_1\sqrt{1-C_1^2\cos^2x-C_2^2\cos^2[\alpha(x-x_0)]}, \tag{17}$$

where

$$\alpha=\frac{\gamma_2}{\gamma_1}=\sqrt{\frac{\beta_2}{\beta_1}},\quad x_0=x_{10}-\frac{\gamma_1}{\gamma_2}x_{20}.$$

Equation (17) has the form of the law of conservation of energy for a mechanical particle,

$$K+U=0,$$

where

$$K=\frac{\dot{x}^2}{2},\quad U=\frac{\gamma_1^2}{2}\left(C_1^2\cos^2 x+C_2^2\cos^2[\alpha(x-x_0)]-1\right).$$

Based on this analogy, we can easily analyze the character of the dynamics of an optical system. The self-consistent propagation of a solitary pulse corresponds to the motion of a mechanical particle from the maximum x_n of potential $U(x)$, determined by the conditions $U'(x_n)=0$ and $U''(x_n)<0$, with zero initial velocity and ending at the maximum x_m. With $x_0=0$ and integral values of α, the maximum points of $U(x)$ are the points $x_n=\pi n$. The initial velocity of a particle from the maximum points equals zero provided that

$$C_1^2+C_2^2=1. \tag{18}$$

In this case, Eq. (17) assumes the form

$$\dot{x}=-\gamma_1\sqrt{C_1^2\sin^2 x+C_2^2\sin^2\alpha x}. \tag{19}$$

(**b**) **Simultons**

The solution of Eq. (19) when $\alpha=1$ has the form

$$\cos x=\tanh\left[\gamma\left(t-\frac{z}{v}\right)\right]. \tag{20}$$

where $\gamma=\gamma_1+\gamma_2$. Consequently, the field amplitudes $A_{1,2}$ and $B_{1,2}$ have the form

$$\begin{aligned} A_1&=\gamma C_1/\cosh\Phi,\quad B_1=C_1\tanh\Phi,\\ A_2&=\gamma C_2/\cosh\Phi,\quad B_2=C_2\tanh\Phi, \end{aligned} \tag{21}$$

where

$$\Phi=\frac{1}{\tau_0}\left(t-\frac{z}{v}\right),\quad \frac{1}{\tau_0}=\gamma=\sqrt{\frac{\beta v}{1-v}}. \tag{22}$$

As can be seen, the solutions given by Eqs. (21) describe the simultaneous propagation of pulses at the adjacent transitions $3\leftrightarrow 2$ and $3\leftrightarrow 1$, appearing when the well-known conditions for the equality of the oscillator strengths at adjacent transitions are satisfied:

$$\beta_1=\beta_2\quad\text{or}\quad \omega_{31}|d_{31}|^2=\omega_{32}|d_{32}|^2, \tag{23}$$

where d_{31} is the matrix element of the dipole moment at the transition $3\leftrightarrow 1$.

In the case $\alpha=2$, the solution of Eq. (19) has the form

$$\begin{aligned} A_1&=\gamma_1 C_1\sqrt{\frac{1+p^2}{\cosh^2\Phi+p^2}},\quad B_1=\frac{C_1\sinh\Phi}{\sqrt{\cosh^2\Phi+p^2}},\\ A_2&=\gamma_2 C_2\,\frac{2\sinh\Phi\sqrt{1+p^2}}{\cosh^2\Phi+p^2},\\ B_2&=C_2\left(1-2\,\frac{1+p^2}{\cosh^2\Phi+p^2}\right), \end{aligned} \tag{24}$$

where

$$\Phi=\frac{1}{\tau_0}\left(t-\frac{z}{v}\right),\quad \frac{1}{\tau_0}=\sqrt{\gamma_1^2C_1^2+\gamma_2^2C_2^2},\quad p=\frac{2C_2}{C_1}. \tag{25}$$

The solutions given by Eqs. (24) are a new form of solutions of the simulton type. The conditions for their appearance differ from the conditions given by Eqs. (23) and have the form

$$4\omega_{31}|d_{31}|^2=\omega_{32}|d_{32}|^2. \tag{26}$$

As can be seen from Eqs. (24), the intensity profiles of the pulses in adjacent transitions are different, which distinguishes the solutions given by Eqs. (24) from the conventional simulton solutions given by Eqs. (21). The intensity profiles of the pulses in adjacent transitions and the dynamics of the population amplitudes of the levels are shown in Fig. 2 for $\alpha=1$ (a and b), $\alpha=2$ (c and d), and $\alpha=3$ (e and f).

(**c**) **Raman solitons**

The solutions given by Eqs. (21) and (24) for Eq. (17) were obtained for $x_0=0$. We now consider the case $x_0=\pi/2\alpha$. Equation (17) in this case acquires the form

$$\dot{x}=-\gamma_1\sqrt{1-C_1^2\cos^2 x-C_2^2\sin^2\alpha x}. \tag{27}$$

For integral α and $C_1>C_2$, potential $U(x)$ reaches maxima at the points $x_n=\pi n$. The initial velocity of a particle from a maximum point equals zero provided that $C_1=1$.

When $\alpha=1$, the solution of Eq. (27) has the form

$$\cos x=\tanh\left[\gamma\sqrt{1-C_2^2}\left(t-\frac{z}{v}\right)\right]. \tag{28}$$

Consequently, the field amplitudes $A_{1,2}$ and $B_{1,2}$ have the form

$$\begin{aligned} A_1&=\gamma/\cosh\Phi,\quad B_1=\tanh\Phi,\\ A_2&=\gamma C_2\tanh\Phi,\quad B_2=C_2/\cosh\Phi, \end{aligned} \tag{29}$$

where

$$\Phi=\frac{1}{\tau_0}\left(t-\frac{z}{v}\right),\quad \frac{1}{\tau_0}=\gamma\sqrt{1-C_2^2}=\sqrt{\frac{\beta v}{1-v}(1-C_2^2)}. \tag{30}$$

As can be seen from Eqs. (29), a ''bright'' soliton occurs in the transition $3\leftrightarrow 1$, and a ''dark'' soliton occurs in the transition $3\leftrightarrow 2$. Such an excitation, by analogy with Refs. 4, 5, 8, and 9, can be called a resonance Raman soliton. The condition for them to be excited is that Eqs. (23) are satisfied.

When $\alpha=2$, the solution of Eq. (27) has the form

$$\begin{aligned} A_1&=\gamma_1 C_1\sqrt{\frac{1-4C_2^2}{\cosh^2\Phi-4C_2^2}},\quad B_1=\frac{C_2\sinh\Phi}{\sqrt{\cosh^2\Phi-4C_2^2}},\\ A_2&=\gamma_2 C_2\,\frac{2\sinh\Phi\sqrt{1-4C_2^2}}{\cosh^2\Phi-4C_2^2}, \end{aligned} \tag{31}$$

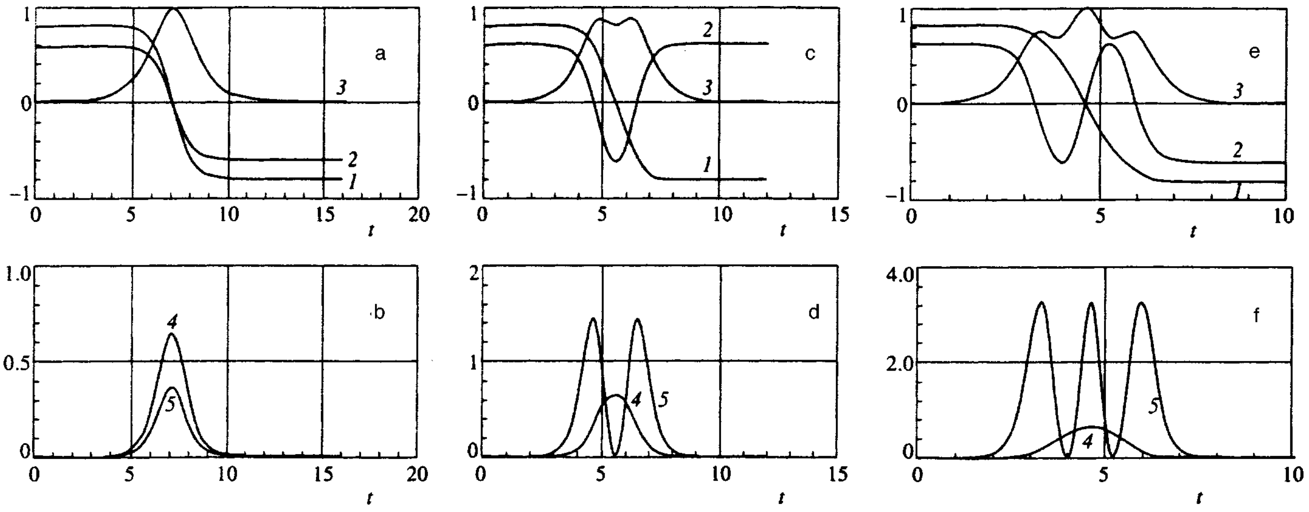


FIG. 2. Dynamics of the population amplitudes of the first (*1*), second (*2*), and third (*3*) levels and intensity profiles of the pulses in the 3↔1 transitions (curve *4*) and the 3↔2 transition (curve *5*) for $\alpha=1$ (a and b), $\alpha=2$ (c and d), and $\omega=3$ (e and f).

$$B_2=C_2\left(1-2\,\frac{1-4C_2^2}{\cosh^2\Phi-4C_2^2}\right),$$

where

$$\Phi=\frac{1}{\tau_0}\left(t-\frac{z}{v}\right),\quad \frac{1}{\tau_0}=\sqrt{\gamma_1^2-\gamma_2^2C_2^2}.$$

It is easy to see that the solutions given by Eqs. (31) are a generalization of the solutions corresponding to resonance Raman solitons for the case $\alpha=2$, i.e., when the frequencies and matrix elements of the transitions satisfy the condition given by Eq. (26).

The intensity profiles of the pulses and the dynamics of the population amplitudes of the levels are shown in Fig. 3 for $\alpha=2$ (a, b, and c) and $\alpha=3$ (d, e, and f).

It can be seen from Eq. (27) that when $C_1=1$ and $\alpha C_2>1$, solutions in the form of a pair of bright and dark solitons exist for an arbitrary value of α. Figure 4 shows the dynamics of the population amplitudes of the levels (a) and the profiles of the bright soliton (b) and the dark soliton (c) for $\alpha=\pi/2$.

(d) The *V* system

As we pointed out above, the relationships between the amplitudes $A_{1,2}$ and $B_{1,2}$ determined by the integral of the motion in Eqs. (11) are qualitatively different in the cases $k_i>0$ and $k_i<0$, which will undoubtedly affect the character of the propagation dynamics of the pulses. We considered the case $k_i>0$ above. Now let $k_i<0$. In this case, the integral of the motion in Eqs. (11) assumes the form

$$B_i^2(z,t)-\frac{1}{\beta_i}\left(1-\frac{1}{v}\right)A_i^2\left(t-\frac{z}{v}\right)=B_{i0}^2-\frac{1}{\beta_i}\left(1-\frac{1}{v}\right)A_{i0}^2. \qquad (32)$$

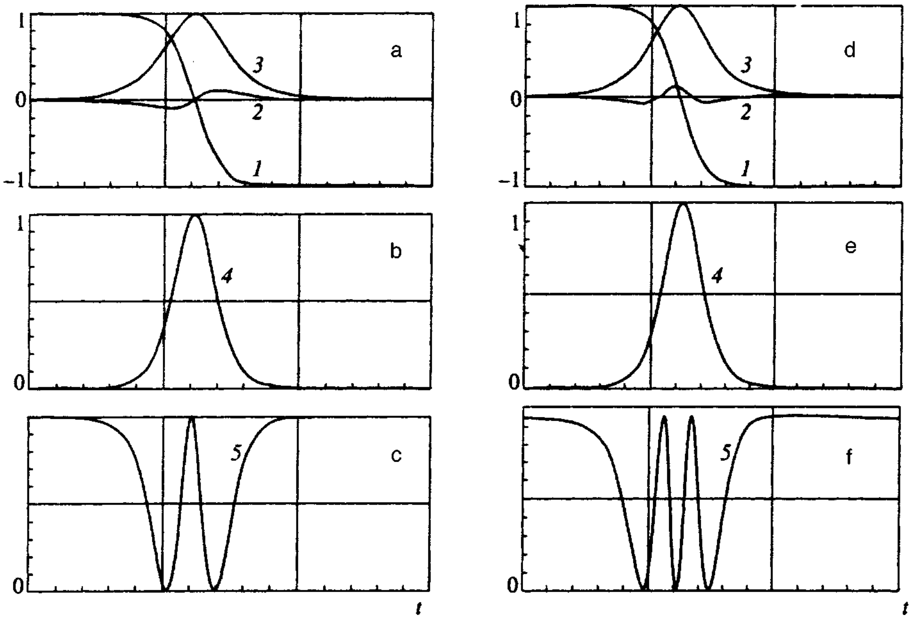


FIG. 3. Dynamics of the population amplitudes of the first (*1*), second (*2*), and third (*3*) levels and intensity profiles of the pulses in the 3↔1 transitions (curve *4*) and the 3↔2 transition (curve *5*) for $\alpha=2$ (a, b, and c), and $\alpha=3$ (d, e, and f).

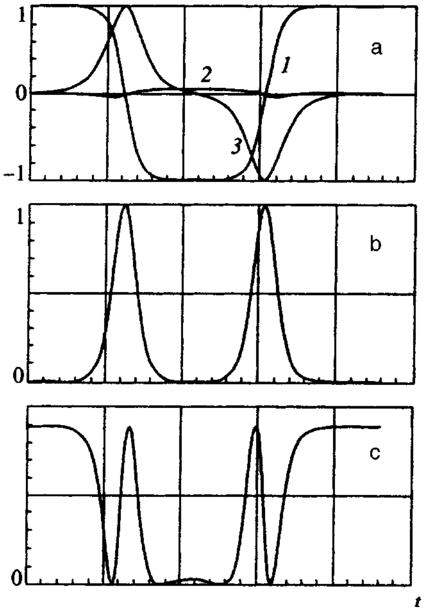


FIG. 4. Dynamics of the population amplitudes of the first (*1*), second (*2*), and third (*3*) levels and intensity profiles of the pulses at the 3↔1 transitions (b) and the 3↔2 transition (c) for the value of the parameter $\alpha=\pi/2$.

Self-consistent solutions in the form of a solitary pulse appear only under the condition that

$$B_{i0}^2-\frac{1}{\beta_i}\left(1-\frac{1}{v}\right)A_{i0}^2=0,$$

which leads to the following connection of amplitudes $A_{1,2}$ and $B_{1,2}$:

$$A_i=\gamma_i B_i, \tag{33}$$

where the parameters γ_i are determined here by the expression

$$\gamma_i=\sqrt{\frac{|\beta_i|v}{1-v}}.$$

Equations (33) lead to the following equations of motion for the amplitudes $B_{1,2}$:

$$\begin{aligned}\dot{B}_1&=\gamma_1 B_1\sqrt{1-B_1^2-B_2^2},\\ \dot{B}_2&=\gamma_2 B_2\sqrt{1-B_1^2-B_2^2}.\end{aligned}\tag{34}$$

It is easy to see that Eqs. (34) have an integral of the motion of the form

$$B_2=CB_1^{\gamma_2/\gamma_1}. \tag{35}$$

Consequently, the coupled motion of the self-consistent pulses in transitions 3↔2 and 3↔1 is described by

$$\dot{B}=\gamma_1 B\sqrt{1-B^2-C^2B^{2\gamma_2/\gamma_1}}. \tag{36}$$

When $\gamma_1=\gamma_2$, the solution of Eq. (36) has the form

$$A_1=\frac{\gamma}{\sqrt{1+C^2}\cosh\Phi},\quad B_1=\frac{1}{\sqrt{1+C^2}\cosh\Phi}, \tag{37}$$

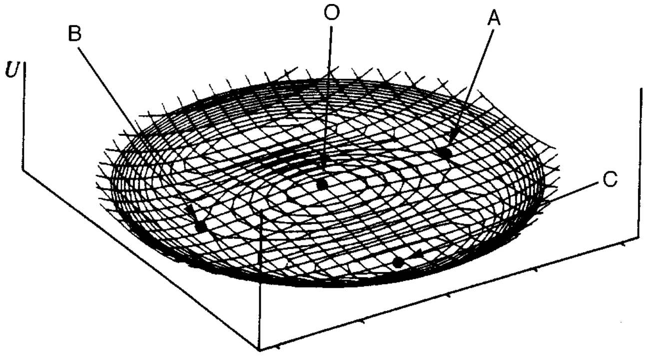


FIG. 5. Form of the potential surface in coordinates (b_1,b_2); *O* is a local maximum point, *A* and *B* are saddle points, and *C* is a local minimum point.

$$A_2=\frac{C\gamma}{\sqrt{1+C^2}\cosh\Phi},\quad B_2=\frac{C}{\sqrt{1+C^2}\cosh\Phi},$$

where

$$\Phi=\frac{1}{\tau_0}\left(t-\frac{z}{v}\right),\quad \frac{1}{\tau_0}=\gamma.$$

As can be seen, these solutions correspond to the case in which all the atoms are in level 3 at the initial instant, which corresponds to bichromatic superradiation for the Λ system, and to copropagating propagation of pulses in resonance with the adjacent transitions for the *V* system.

(e) Coherent trapping of populations

As we pointed out at the beginning of this section, the case in which the inequality

$$|b_3(t\to\pm\infty)|<|b_2(t\to\pm\infty)|<|b_1(t\to\pm\infty)|$$

is not satisfied corresponds to an initially excited medium. In this case, the equations of motion for the variables $b_{1,2}$ can result in a Hamiltonian form with potential energy

$$U(b_1,b_2)=\frac{\gamma_1}{2}(b_1^4-b_1^2)+\frac{\gamma_2}{2}(b_2^4-b_2^2)+\frac{\gamma_1+\gamma_2}{2}b_1^2b_2^2.$$

This potential has the form of the bowl shown in Fig. 5. The potential has the following extrema: $b_1=b_2=0$ (point *O*), $b_1=0$, $b_2=1/2$ (point *A*), $b_1=1/2$, $b_2=0$ (point *C*) and the points symmetric to them. Point *O* is the maximum, while points *A* and *C*, depending on the ratio of the parameters γ_1 and γ_2, are either saddle points or minima. Solutions that start from a saddle point correspond to Raman solitons, while the minimum of the potential corresponds to the state of coherent trapping of populations.

5. THE NONRESONANCE CASE

(a) The *V* system

In the nonresonance case, along with the integral of the motion given by Eqs. (11), we must also take into account the integral of the motion given by Eqs. (7). The combined solution of these two equations determines the conditions for the coupling of amplitudes $A_{1,2}$ and $B_{1,2}$:

$$\begin{gathered}(\dot{\varphi}_i+\varphi_i')A_i^2+\beta_i\dot{\psi}_iB_i^2=0,\\ B_i^2+\frac{1}{\beta_i}\left(\frac{1}{v}-1\right)A_i^2=B_{i0}^2+\frac{1}{\beta_i}\left(\frac{1}{v}-1\right)A_{i0}^2.\end{gathered}\tag{38}$$

Comparison of these two equations shows that when $k_i>0$, a self-consistent solution can be only phase-modulated. Indeed, we get from Eqs. (38) that

$$B_i^2\left[\frac{\dot{\varphi}_i+\varphi_i'+(1-1/v)\dot{\psi}_i}{\dot{\varphi}_i+\varphi_i'}\right]=C_i^2.$$

Consequently, the B_i are constants for constant $\dot{\varphi}_i$, $\dot{\psi}_i$, and φ_i'.

In the case $k_i<0$, the quantities $\dot{\varphi}_i$, $\dot{\psi}_i$, and φ_i' can be constant; we obtain from Eqs. (38) that

$$\frac{1}{\beta_i}\left(1-\frac{1}{v}\right)=-\frac{\dot{\varphi}_i+\varphi_i'}{\beta_i\dot{\psi}_i}.$$

Assuming that

$$\varphi_i(z,t)=\varphi_i\left(t-\frac{z}{v}\right),\tag{39}$$

we get

$$\dot{\varphi}_i=\Delta_i,\quad \dot{\psi}_i=-\Delta_i.$$

Thus, in the case $k_i>0$, the equations of motion, Eqs. (34), are easy to generalize and take the form

$$\begin{gathered}\dot{B}_1=\pm\gamma_1B_1\sqrt{1-\left(\frac{\Delta_1}{\gamma_1}\right)^2-B_1^2-B_2^2},\\ \dot{B}_2=\pm\gamma_2B_2\sqrt{1-\left(\frac{\Delta_2}{\gamma_2}\right)^2-B_1^2-B_2^2}.\end{gathered}\tag{40}$$

When the condition

$$\Delta_1=\Delta_2\sqrt{\beta_1/\beta_2},\tag{41}$$

is satisfied, the system of Eqs. (40) again has an integral of the motion of the form of Eq. (35), and the equation of coupled self-consistent motion of the pulses takes the form

$$\dot{B}=\pm\gamma_1B\sqrt{1-(\Delta_1/\gamma_1)^2-B^2-(1+C^2)B^{2\gamma_2/\gamma_1}}.\tag{42}$$

The solution of Eq. (42) in the case $\gamma_1=\gamma_2$ has the form

$$\begin{gathered}A_1=\sqrt{\frac{1-(\Delta/\gamma)^2}{1+C^2}}\frac{\gamma}{\cosh\Phi},\\ B_1=\sqrt{\frac{1-(\Delta/\gamma)^2}{1+C^2}}\frac{1}{\cosh\Phi},\\ A_2=\sqrt{\frac{1-(\Delta/\gamma)^2}{1+C^2}}\frac{\gamma C}{\cosh\Phi},\\ B_2=\sqrt{\frac{1-(\Delta/\gamma)^2}{1+C^2}}\frac{C}{\cosh\Phi},\end{gathered}\tag{43}$$

where

$$\Phi=\frac{1}{\tau_0}\left(t-\frac{z}{v}\right),\quad \frac{1}{\tau_0}=\sqrt{\gamma^2-\Delta^2}.$$

(b) Phase-modulated simultons

In the case $k_i>0$, a self-consistent solution can have the form of only a phase-modulated pulse. As follows from the first equation of the system of Eqs. (38), $\dot{\varphi}_i$ and $\dot{\psi}_i$ can be represented in the form

$$\dot{\varphi}_i=\kappa_iB_i^2,\quad \dot{\psi}_i=\mu_iA_i^2,\tag{44}$$

where

$$\mu_i=\frac{\kappa_i}{\beta_i}\left(\frac{1}{v}-1\right)=\frac{\kappa_i}{\gamma_i^2}.$$

Substituting Eq. (44) into the second Eq. (9), it is easy to get the following equations for the amplitudes $B_{1,2}$:

$$\begin{gathered}\dot{B}_1=\gamma_1\sqrt{C_1^2-B_1^2}\sqrt{1-(1+C_1^2\Gamma_1^2)B_1^2+\Gamma_1^2B_1^4-B_2^2},\\ \dot{B}_2=\gamma_2\sqrt{C_2^2-B_2^2}\sqrt{1-(1+C_2^2\Gamma_2^2)B_2^2+\Gamma_2^2B_2^4-B_1^2}.\end{gathered}\tag{45}$$

where C_i is determined by Eq. (13), while

$$\Gamma_i=\mu_i\gamma_i=\kappa_i\Big/\sqrt{\frac{\beta_iv}{1-v}}.$$

Let us introduce the following replacement of variables:

$$B_i=C_i\cos\gamma_ix_i,\quad A_i=\gamma_iC_i\sin\gamma_ix_i,$$

and then the system of Eqs. (45) assumes the form

$$\begin{gathered}\dot{x}_1=-\sqrt{1-C_1^2\cos^2\gamma_1x_1-C_2^2\cos^2\gamma_2x_2-\Gamma_1^2C_1^4\sin\gamma_1x_1\cos\gamma_1x_1},\\ \dot{x}_2=-\sqrt{1-C_1^2\cos^2\gamma_1x_1-C_2^2\cos^2\gamma_2x_2-\Gamma_2^2C_2^4\sin\gamma_2x_2\cos\gamma_2x_2}.\end{gathered}\tag{46}$$

It can be seen from Eqs. (46) that when $C_1^2+C_2^2=1$, the velocities are $\dot{x}_{1,2}=0$ at the points

$$\gamma_1x_{1n}=\pi n,\quad \gamma_2x_{2m}=\pi m.$$

Then, in the case $C_1^2+C_2^2=1$ and $\Gamma_1C_1=\Gamma_2C_2$, the solution of the system of Eqs. (45) has the form

$$\begin{gathered}A_1=\gamma_1C_1\sqrt{\frac{1-\Gamma_1^2C_1^2}{\cosh^2\Phi-\Gamma_1^2C_1^2}},\quad B_1=\frac{C_1\sinh\Phi}{\sqrt{\cosh^2\Phi-\Gamma_1^2C_1^2}},\\ A_2=\frac{\kappa_1}{\kappa_2}A_1,\quad B_2=\frac{\kappa_1}{\kappa_2}B_1\end{gathered}\tag{47}$$

and

$$\dot{\varphi}_1=\kappa_1\,\frac{C_1^2\,\sinh^2\,\Phi}{\cosh^2\,\Phi-\Gamma_1^2C_1^2},$$

where

$$\Phi=\frac{1}{\tau_0}\left(t-\frac{z}{v}\right),\quad \frac{1}{\tau_0}=C_1\sqrt{\gamma_1^2-C_1^2\kappa_1^2}. \tag{48}$$

(c) Phase-modulated Raman soliton

When $C_1=1$, the points at which $\dot{B}_{1,2}=0$ will be $B_1=1$, $B_2=C_2$. Using this circumstance, it is easy to obtain coupled self-consistent solutions of the system of Eqs. (48) in the form of solitary pulses. Thus, when $\gamma_1=\gamma_2$ and $\Gamma_1=\Gamma_2C_2^2$, they have the form

$$\begin{aligned} A_1&=\gamma_1\sqrt{\frac{1-\Gamma_1^2}{\cosh^2\,\Phi-\Gamma_1^2}},\quad B_1=\frac{\sinh\,\Phi}{\sqrt{\cosh^2\,\Phi-\Gamma_1^2}},\\ A_1&=\frac{C_2\,\sinh\,\Phi}{\sqrt{\cosh^2\,\Phi-\Gamma_1^2}},\quad B_1=C_2\sqrt{\frac{1-\Gamma_1^2}{\cosh^2\,\Phi-\Gamma_1^2}}. \end{aligned} \tag{49}$$

The dynamics of the phases of the pulses are determined by the equations

$$\dot{\varphi}_1=\kappa_1\,\frac{\sinh^2\,\Phi}{\cosh^2\,\Phi-\Gamma_1^2},\quad \dot{\varphi}_2=\kappa_2\,\frac{C_2^2(1-\Gamma_1^2)}{\cosh^2\,\Phi-\Gamma_1^2}, \tag{50}$$

where

$$\Phi=\frac{1}{\tau_0}\left(t-\frac{z}{v}\right),\quad \frac{1}{\tau_0}=\gamma_1\sqrt{1-\Gamma_1^2-C_2^2}.$$

It can be seen from Eqs. (50) that as $t\to\pm\infty$, the pump pulse, i.e., the pulse in quasi-resonance with transition $1\leftrightarrow 3$, is offset from exact resonance by κ_1. At the maximum of the pump pulse, its offset from resonance goes to zero, whereas the offset of the Stokes pulse reaches a maximum equal to $\dot{\varphi}_2=\kappa_2C_2^2$ at that instant.

6. CONCLUSION

The analysis given here has made it possible to determine the form of self-consistent solutions in the form of steady-state pulses in adjacent transitions of a three-level atom whose intensity profile has both a unimodal and a nonunimodal shape. The conditions for steady-state pulses to appear depend upon the parameter α, which is determined by the ratio of the oscillator strengths on the adjacent transitions:

$$\alpha=\sqrt{\frac{\beta_2}{\beta_1}}=\sqrt{\frac{\omega_{32}|d_{32}|^2}{\omega_{31}|d_{31}|^2}},$$

the parameters $C_{1,2}$, which are determined by the initial populations of the levels and the field amplitudes, and the parameters κ_i, which determine the initial offsets of the pulse frequencies or the value of the chirp.

In exact resonance, simulton solutions for the Λ system arise for integral values of α. In this case, a pulse at a transition with a large oscillator strength has multiple peaks, the number of peaks being α. Solutions in the form of Raman solitons appear for an arbitrary value of α, and the number of dips in a dark soliton is determined by the integer part of the number α. In the case of the V system, simultons can exist both in the resonance case and in the nonresonance case. In the nonresonance case, the frequency offsets of the pulses at adjacent transitions are related by Eq. (41). Solutions in the form of a phase-modulated simulton and Raman soliton arise for certain relationships (indicated by us) between the initial offsets of the pulses and the oscillator strengths at adjacent transitions, and depend on the initial populations and field amplitudes.

Since the profiles of the steady-state pulses are described in analytic form, this can be extremely useful in studying the properties of soliton-like excitations observed in various experiments.

The author is grateful to S. R. Durlevich for help in preparing the illustrations for this paper.

This work was carried out with the financial support of the Russian Fund for Fundamental Research (Project 96-02-19285).

*)e-mail: andreev@srl.ilc.msu.su

---

[1] L. A. Bol'shov and V. V. Likhanskiĭ, Kvant. Elektron. (Moscow) **12**, 1339 (1985) [Sov. J. Quantum Electron. **15**, 889 (1985)].

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

[3] B. D. Agap'ev, M. B. Gornyĭ, B. G. Matisov, and Yu. V. Rozhdestvenskiĭ, Usp. Fiz. Nauk **163**, 1 (1993) [Sov. Phys. Usp. **36**, 763 (1993)].

[4] K. Drühl, R. G. Wenzel, and J. L. Carlsten, Phys. Rev. Lett. **51**, 1171 (1983).

[5] D. C. MacPherson, R. C. Swanson, and J. L. Carlsten, Phys. Rev. A **40**, 6745 (1989).

[6] D. E. Gakhovich, A. S. Grabchikov, and V. A. Orlovich, Opt. Commun. **102**, 485 (1993).

[7] D. P. Scherrer, M. Baumgartner, and F. K. Kneubühl, in *Proceedings of the Twelfth International Conference on Infrared and Millimeter Waves*, Orlando, Florida, Dec., 1995, p. 162.

[8] J. C. Englund and C. M. Bowden, Phys. Rev. A **46**, 578 (1992).

[9] C. Claude, F. Ginovart, and J. Leon, Physica **87**, 348 (1995).

[10] A. I. Maimistov, A. M. Basharov, S. O. Elyutin, and Yu. M. Skylarov, Phys. Rep. **191**, 1 (1990).

[11] A. V. Andreev, V. I. Emel'yanov, and Yu. A. Il'inskiĭ, *Cooperative Phenomena in Optics*, Nauka, Moscow (1988).

Translated by W. J. Manthey

ERRATA

Erratum: Spectral properties of backward stimulated scattering in liquid carbon disulfide [JETP 85, 850–856 (November 1997)]

Guang S. He, Yiping Cui, and Paras N. Prasad

Photonics Research Laboratory, State University of New York at Buffalo, Buffalo, New York 14260-3000, USA

[S1063-7761(98)02902-3]

The correct caption of Fig. 4 should read as ''... at various attenuation ratios: 1/180 (a), ...''
The correct caption of Fig. 5 should read as ''pumped by the ~0.1 cm^{-1}-wide 532-nm line...''
The correct Ref. 11 should be: E. J. Miller and R. W. Boyd, Int. J. Nonlinear Opt. Phys. **4**, 765 (1992).