ISSN (print) 0044-4510
ISSN (online) 3034-641X

Российская Академия Наук

Volume 167, No. 1 January 2025

ЖУРНАЛ ЭКСПЕРИМЕНТАЛЬНОЙ *и* ТЕОРЕТИЧЕСКОЙ ФИЗИКИ

JOURNAL OF EXPERIMENTAL *and* THEORETICAL PHYSICS

NAUKA
1727

This Journal includes translation of articles published
in *Журнал экспериментальной и теоретической физики / Journal of Experimental and Theoretical Physics (Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki), Vol. 167, No 01 (2025)*

Founders:

Russian Academy of Sciences,
P.L. Kapitza Institute for Physical Problems
of the Russian Academy of Sciences

Translated by Nauka Publishing House
under the Contract No. 4U-EP-020-2-25 dated March 11, 2025

RUSSIAN ACADEMY OF SCIENCES

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS

FOUNDED IN MARCH 1873
PUBLISHED 12 TIMES A YEAR
MOSCOW

VOLUME 167, ISSUE 1
JANUARY 2025
RAS

THE JOURNAL IS PUBLISHED UNDER THE SUPERVISION
OF THE PHYSICAL SCIENCES DEPARTMENT OF THE RUSSIAN ACADEMY OF SCIENCES

CONTENTS

ATOMS, MOLECULES, OPTICS

NUCLEI, PARTICLES, FIELDS, GRAVITATION, AND ASTROPHYSICS

SOLIDS AND LIQUIDS

ORDER, DISORDER, AND PHASE TRANSITION IN CONDENSED MEDIA

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 3–23

ATOMS, MOLECULES, OPTICS

ON THE LIFETIME OF QUASI-STATIONARY LEVELS DURING TUNNELING IN A RESONANT TUNNEL STRUCTURE

© 2025 M.V. Davidovich[a,*], I.S. Nefedov[a, b]

[a] National Research Saratov State University named after N.G. Chernyshevsky, Saratov, Russia
[b] RUDN University, Moscow, Russia
*e-mail: davidovichmv@yandex.ru

Received March 08, 2024

Revised March 08, 2024

Accepted March 28, 2024

Abstract. We investigate field emission in resonant tunneling heterostructures with one quantum well and two barriers, as well as the influence of the lifetime of resonant metastable levels formed in the well on it. The problem of the tunneling time of a quantum particle (electron) through a structure with a barrier and two barriers and a well is also considered. Stationary and nonstationary Schrodinger equations are used. The lifetimes of metastable levels are determined and their effect on tunnel current is investigated.

DOI: 10.31857/S00444510250101e9

1. INTRODUCTION

Resonant tunneling structures (RTS), which are nanoscale heterostructures with field emission, are widely used in electronics as sources of very high current densities (up to 10^{14} A/m²) [1–3]. They also form the basis for designing resonant tunneling diodes (RTD), transistors, quantum cascade lasers (QCL) of the "Stark ladder" type, THz transistors and switches [4–11], as well as other devices. For all these applications, switching times, response times, and overall transient times are crucial [11,12].

RTS is characterized by having one, two, or several quantum wells separated by barriers, where quasi-stationary resonance levels can arise [13]. We will refer to the electron-emitting left electrode as the cathode (denoted by the subscript c) and the right electrode as the anode (subscript a). For the intermediate electrode (grid), the subscript g will be used. In the case of equal electrochemical potentials of the electrodes $\mu_c = \mu_a$, the quantum potential V(x) between the two electrodes (in a diode structure) resembles an inverted parabola on a pedestal and can be strictly described by an infinite series of images that account for the electron work functions [1]. Approximating this series with an inverted parabola is rather crude [1,14]. A more accurate approximation is a fourth-order inverted parabola [1,2]. In this work, we will use an even more precise approximation for the potential in the diode $0 < x < d$ under anode voltage U_a:

$$V(x) = E_{Fc} + W_c \frac{(1-\alpha/d)(1+\delta/d)^2}{(1-\delta/d)^2\,\varepsilon} \times$$

$$\times\left[1 - \frac{\delta d}{\big(x+\delta(1-x/d)\big)\big(d-x+x\delta/d\big)}\right] - \frac{eU_a x}{d}. \quad (1)$$

In this formula, $\alpha = \delta\big(2\ln(2)+1\big)$ represents the cathode work function, assumed equal for both the cathode and anode $W_c = W_a$, and is related to the parameter (gap size) δ by the equation:

$$W_a = e^2/(16\pi\varepsilon_0\delta).$$

For simplicity, we will further assume equal Fermi energies $E_{Fc} = E_{Fa}$ for the electrodes. Under potential $V(x)V(x)V(x)$, the boundary conditions are $V(0) = E_{Fa}$ at the cathode and $V(d) = E_{Fa} - eU_a = \mu_a$, at the anode, meaning the quantum potential V coincides with the electrochemical potentials. In cases with different work functions (and materials of the cathode and anode), an additional term $(E_{Fa} - E_{Fc})x/d$ should be added to (1). The accuracy of equation (1) is no worse than 1%. Diode structures do not allow for extremely high current densities. Current increase

occurs through resonant tunneling (RT), where one or more quantum wells are surrounded by barriers [1–3]. Reflections from the barriers interfere, and the total reflection coefficient cancels out. For simplicity, we further consider a single-well structure with three electrodes: cathode (source), grid (gate), and anode (drain). Equation (1) applies both to a vacuum gap ($\varepsilon = 1$), and a dielectric gap between electrodes. In the absence of an anode voltage $U_a = 0$, the potential in the center of the gap is:

$$V\left(d/2\right) = E_{Fc} + \frac{W_c\left(1-\alpha/d\right)\left(1+\delta/d\right)^2}{\left(1-\delta/d\right)^2\varepsilon}.$$

For a work function of 3.6 eV, the corresponding value is $\delta = 0.1$. Thus, for typical work functions of materials (2–5 eV) and typical electrode and gap sizes, RTS structures on the order of nanometers satisfy the inequalities $\delta/d << 1$, $\alpha / d << 1$. In the absence of anode voltage, the inequality $V\left(d/2\right) \approx E_{Fc} + W_c/\varepsilon$ holds. A dielectric with dielectric permittivity ε reduces the barrier height by a factor of ε.

Suitable and convenient dielectrics for RTS include CVD (Chemical Vapor Deposition) diamond ($\varepsilon = 5.6$, bandgap 2.5 eV) [15] and beryllium oxide (BeO, $\varepsilon = 6.7$, bandgap 10.6 eV). These dielectrics significantly reduce the barrier height and have the highest thermal conductivity, which is essential for high current densities [2,3]. Although CVD diamond with 88% sp^3 hybridization has a density of 88.2% of crystalline diamond, its dielectric constant can be taken as 5.6 due to the presence of a small graphite phase. Electrodes can be made of metals or doped semiconductors. Beryllium has the highest Fermi energy (14.6 eV), relatively low work function (3.92 eV), and the highest thermal conductivity among metals. To construct a complex profile V (Fig. 1), equation (1) is applied twice – once for the cathode-grid gap (replacing $U_a \to U_g$) and once for the grid-anode gap, assuming $E_{Fc} \to E_{Fc} - eU_g$. On the grid, the quantum potential is constant and determined by its electrostatic potential U_g.

In RTDs and QCLs, highly conductive layers are usually considered electrically free, meaning the potential along them is not fixed and decreases [4–12]. Figure 1 shows typical profiles of V(x) for diode and triode structures under different anode (U_a) and grid (U_g) voltages for copper electrodes. To form a quantum well, a grid voltage $U_g = E_{Fc}/e$ was applied. The energy E is measured from the

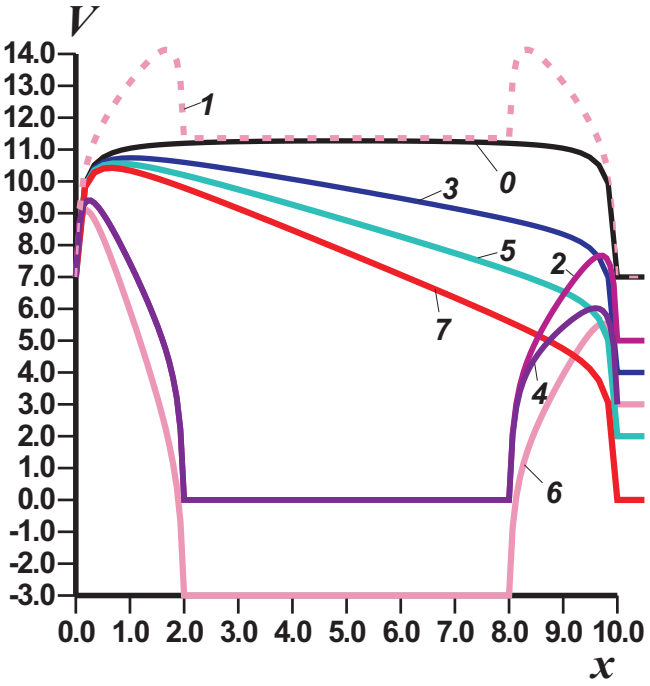

Fig. 1. Potential Barrier Profile V (eV) as a Function of Distance x (nm) in a vacuum diode (curves 0, 3, 5, 7) and a vacuum triode (curves 1, 2, 4, 6). The curve numbers for the diode correspond to the anode voltage U_a in volts. For the triode curves 2 and 4, the grid voltage $U_g = E_F$ is specified, and their numbers correspond to the anode voltage. For curve 1, $U = 0$, $U_g = W_c/e$, while for curve 6, $U_a = 4$ V and $U_g = 3$ V. The work functions are $E_{F_c} = 7$ eV and $W_c = 4.36$ eV (copper electrodes).

conduction band bottom of the cathode, which coincides with the bottom of the well.

If an energy level E_n exists in the formed quantum well, it is quasi-stationary, as there is always an identical level at both the cathode and anode, allowing the electron to tunnel between them. Tunneling can occur both leftward to the cathode and rightward to the anode. Subsequently, the electron transitions from this level to the Fermi level of the corresponding electrode, from which it can enter the power supply circuit, as only electrons near the Fermi level participate in the diffusion current.

The issue of quasi-stationary level lifetime (decay time) in a spherically symmetric quantum well has been addressed in several works, such as [16–18]. However, the lifetime of quasi-stationary levels in a one-dimensional Cartesian RTS has not been strictly studied. This time is closely related to the tunneling time of a single particle through the structure (its dwell time). There is extensive literature on the introduction of various time definitions (see, for example, the reference list in [19]). The topic

of tunneling times remains under discussion, with several paradoxes. Tunneling times are closely linked to the switching times of tunnel devices [11]. RTS devices such as RTDs, QCLs, and other structures are typically modeled using rectangular potentials modified by the term $-eU_a x/d$ [11]. This approach is a rough approximation because high voltages lead to a barrier shape close to a triangle on a pedestal [1]. Schrödinger equation (SE) calculations show that such a barrier is orders of magnitude more transparent than a rectangular barrier of the same height and base width. A semiclassical approximation is often used [16], which can be integrated exactly for a triangular barrier [20]. However, this method is accurate only up to a pre-exponential factor and is quite imprecise in the narrow upper part of the barrier, as it neglects the reflected electron wave [20]. For narrow barriers, the reflected wave contributes significantly. In the analysis of RTS with two or more rectangular barriers, resonance levels are usually defined as the penetration of a particle through identical barriers to the left or right with the same energy as in the well [13]. However, real RTS structures differ. Upon reaching the cathode or anode with a given energy, the particle transitions to the Fermi level of the electrode, emitting or absorbing an energy quantum, and exits the structure with this energy, as any current in conductors is generated by electrons near the Fermi level. Under stationary tunneling (constant anode voltage), the number of electrons tunneling from the cathode is exponentially greater than the number tunneling from the anode, resulting in a constant emission current closing through the power supply. The emergence of resonance levels E'_n leads to resonant tunneling (RT), accompanied by an increase in current, as the barrier becomes fully transparent for electrons with energy $E = E'_n$. Quasi-stationary levels arise with increasing well width. These energy levels are complex: $\tilde{E}_n = E'_n - iE''_n$. The parameter E''_n determines the level lifetime $\tau_n = 2\hbar / E''_n$. The smaller the lifetime, the broader the energy level, the wider the energy range satisfying the condition $E \approx E'_n$, and the greater the number of electrons undergoing resonant tunneling. Thus, determining the lifetimes (complex energies $\tilde{E}_n$) of quasi-stationary levels and their dependence on quantum potential configurations is crucial, which is the primary aim of this study. For field emission, the number of electrons incident per second on the barrier within a velocity interval $v_z + dv_z$ and energy range is: $d\nu(v_z) = n^+(k)v_z dv_z$,

where

$$n^+(k) = \frac{m_e^2 k_B T}{2\pi^2\hbar^3} \times \times \ln\left(1 + \exp\left(\frac{E_{Fc} - E(k)}{k_B T}\right)\right) v_z. \quad (2)$$

Equation (2) is derived by averaging over all transverse velocities of the Fermi gas electrons in the metal cathode and is presented for finite temperatures. For cold emission (T = 0), the spectrum is limited by the Fermi energy:

$$n^+(k) = m_e^2\left(E_{Fc} - E(k)\right)/\left(2\pi^2\hbar^3\right)v_z.$$

Although the actual tunneling process involves a multi-speed electron flux determining the total tunnel current density:

$$J^+(U_a) = \frac{-em_e}{2\pi^2\hbar^3}\int_0^{\mu_c} D^+(E,U_a)(\mu_c - E)dE, \quad (3)$$

this problem can be treated as single-particle tunneling with a specified energy E.

The electron charge is taken as $q_e = -e$, so the positive electron flux from the cathode results in a positive anode current $-J^+(U_a)$ through a unit cross-section. The upper limit in equation (3) is on the order of several electronvolts, which is consistent with non-relativistic quantum mechanics. For thermionic-field emission (at T~2000K), equation (2) should be used, with the upper limit in equation (3) extended by a few eV due to the logarithmic decay.

For T = 0, the total current density $J = J^+ - J^-$ is determined by tunneling in both directions with transmission coefficients $D^\pm(E) = 1 - \left|R^\pm\right|^2$, derived from reflection coefficients $R^\pm$. To determine $R^\pm$, the Schrödinger equation is solved. The expression for J^- is obtained by substituting $\mu_c \to \mu_a$, $D^+ \to D^-$.

For a symmetric potential ($U_a = 0$), the tunneling coefficient T(E)T(E)T(E) is always $D^+ = D^-$. For a weakly asymmetric potentials.

2. LIFETIME OF THE LEVEL BASED ON THE STATIONARY SCHRÖDINGER EQUATION SOLUTION

The stationary Schrödinger equation (SSE):

$$\left(-\frac{(\hbar\partial_x)^2}{2m_e} + V(x)\right)\psi(x) = 0$$

is most conveniently solved for V(x) using the wave impedance transformation method. For a constant potential V_n in the region $x_n < x < x_{n+1}$, the wave impedance is introduced as:

$$z_n(E) = -i\psi(x) / \psi'(x) = 1 / k_n,$$

where:

$$\psi(x) = A \exp(ik_n x)$$

is the wave function (WF) of an electron moving in the direction of x electron,

$$k_n = \sqrt{2me(E - V_n)}/\hbar.$$

Let $z_0(x_{n+1})$ be the impedance on the right side. It transforms into the input impedance on the left side according to the formula:

$$Z_i(x_n) = z_n \frac{z_0 - iz_n \mathrm{tg}(k_n(x_{n+1} - x_n))}{z_n - iz_0 \mathrm{tg}(k_n(x_{n+1} - x_n))}. \quad (4)$$

Setting $z_0 = Z_i(x_n)$, we apply this formula iteratively for each segment until we obtain the input impedance at the cathode $Z_{ic}(0)$ and the reflection coefficient from the cathode side:

$$R^+ = (1 - k_0 Z_{ic}(0)) / (1 + k_0 Z_{ic}(0)).$$

Here,

$$k_0 = k_c = \sqrt{2m_e E} / \hbar.$$

For the initial iteration at the anode, we assume:

$$k_a = \sqrt{2m_e(E - E_{Fc} + eU_a)}/\hbar,$$
$$z_0 = 1/k_a.$$

It is worth noting that in typical tunneling through a barrier, $k_a = k_0$ is taken, i.e., the motion is considered only up to the turning point. Such transparency is $\tilde{D}^+ = \tilde{D}^-$. However, after passing this point, the electron moves quasi-classically, gaining energy eU_a. This results in lowering the Fermi level at the anode by eU_a, necessitating the use of the adjusted value k_a.

This concept can be illustrated using an infinitely narrow step-like barrier: $V = 0$ at $x < 0$ and $V = -eU_a$ at $x > 0$. For such a barrier, the quasi-classical approximation gives full transparency, $D = 1$, $R = 0$. However, under the strict solution, the reflection from the step is:

$$R = (k_0 - k_a) / (k_0 + k_a)$$

and $D < 1$. Applying formula (4) is equivalent to matching the wave function and its derivative. Clearly, the energy levels $E_n = E_n' - iE_n''$ can be defined as the complex roots of the equation:

$$R^+(E_n) = 0.$$

The transparency from the anode to the cathode D^- is determined by reverse transformation, where at the anode we take:

$$k_a = \sqrt{2m_e E} / \hbar,$$
$$z_0 = 1 / k_a$$

and

$$R^- = (1 - k_0 Z_{ia}(d)) / (1 + k_0 Z_{ia}(d)).$$

The difference between D^+ and D^- increases with increasing U_a. When $eU_a > E_{Fc}$, tunneling from the anode becomes impossible. After tunneling, the electron always transitions to the Fermi level of the corresponding electrode, either releasing or absorbing energy $e|E - E_{F(a,c)}|$ depending on the sign of the energy difference. This process is diffusive, occurring over a distance on the order of the electron mean free path, and does not affect the wave tunneling process itself. If tunneling occurs from a level below the Fermi energy, heating of the corresponding electrode occurs (Nottingham effect): the departing electron is replaced by an electron from the Fermi level. For $U_a = 0$, we obtain a symmetric structure in the form of a quantum well between two barriers (see Fig. 1, curves 0 and 1). In this case, the condition:

$$R^+(E_n) = R^-(E_n) = 0$$

yields energy levels from which the particle can tunnel equally to the left or right. Otherwise, the condition $R^+(E_n) = 0$ gives the levels from which the particle can escape to the anode, while $R^-(E_n) = 0$ corresponds to levels leading to cathode transitions. Calculations show that the levels approximately coincide within their width. For example, if $eU_a > E_{FA}$, all energy levels at the anode become negative, making transitions to positive energy levels on the cathode impossible. It is evident that for $E < 0$, when $|R^-(E)| \equiv 1$, i.e. meaning no solutions exist for the equation $R^-(E) = 0$. In this case, the cathode impedance $1/k_c$ becomes imaginary, and the cathode acts as an infinitely long, fully reflective

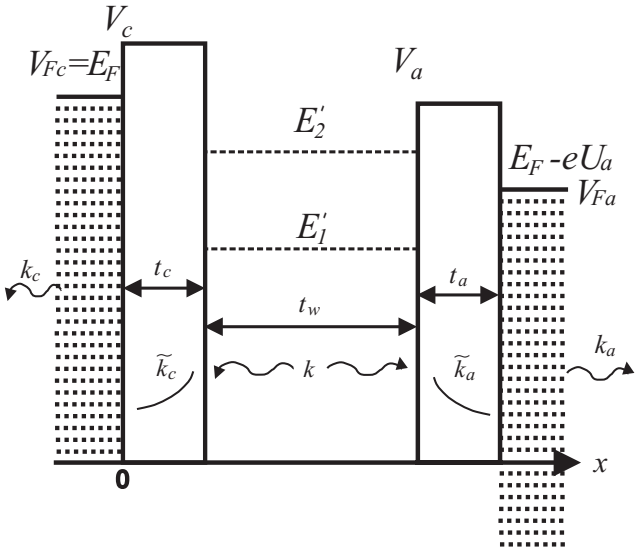

Fig. 2. Schematic potential distribution V in a single-well RTS at $U_g = E_F/e$. Dashed lines indicate the energy levels at the cathode, anode, and two metastable levels.

step for the anode. Positive energies at the anode can only exist at non-zero temperatures, i.e., under thermionic emission conditions. Solutions to the equation $R^+(E) = 0$ always exist for levels on the cathode side. Thus, for an asymmetric potential, two types of energy levels exist. Resonant tunneling is primarily considered for asymmetric potentials, as this condition ensures a continuous current.

Another possible approach to solving the stationary Schrödinger equation involves using transfer matrices [1–3] $\hat{T}(E)$. The structure matrix is defined by piecewise-constant potential V approximations and multiplying the segment matrices. The characteristic equation for determining tunneling levels at the anode takes the form [1, 3]:

$$ik_c(E) = \frac{T_{21}(E) - ik_a(E)T_{22}(E)}{T_{11}(E) - ik_a(E)T_{12}(E)}. \qquad (5)$$

Another method involves using the sweep method. In addition to finding $R^{\pm}(E)$ and $D^{\pm}(E)$ this approach allows for determining the wave function amplitudes $A_n^{\pm}\psi\left(\pm ik_n(x - x_n)\right)$ and the charge distribution in the barrier and well region under known incident particle fluxes from the cathode $n^+(k)v_z dv_z$ and the anode $n^-(k)v_z dv_z$.

This, in turn, enables the estimation of changes in the quantum potential V due to space charge effects under high currents [2]. Such estimation requires iterative solutions of the Poisson equation (PE) and the Schrödinger equation. However, these numerical methods are less convenient for our analysis of resonant level influence on electron emission.

We derive the exact solution of the Schrödinger equation (SE) for the model potential $\tilde{V}(x)$, described by two rectangular barriers of height V_c at the cathode and V_a at the anode (see Fig. 2). To better match the real potential, the barrier widths t_c and T_a are taken approximately half the size of the bases of the actual near-triangular barriers on a rectangular pedestal (Fig. 1, curves 2, 4, 6), while the well width t_u is correspondingly increased. It is possible to achieve an exact correspondence between the width of a triangular barrier and the width of a rectangular barrier with equal heights by equating their transparencies $D_{rec}(E) = D_{tre}(E)$. This correspondence depends on the energy. Averaging over the energy range, we obtain a coefficient of approximately $t_{rec} \approx 0.5t_{tre}$. In the quantum well, the SE solution takes the form:

$$\psi(x) = A_w^+ \exp\left(ik_0(x - t_c)\right) + A_w^- \exp\left(-ik_0(x - t_c)\right).$$

In the barrier region near the cathode, the wave function (WF) is:

$$\psi(x) = A_c^+ \exp\left(-\tilde{k}_A x\right) + A_c^- \exp\left(\tilde{k}_A x\right),$$

Similarly, in the barrier region near the anode, the wave function is:

$$\psi(x) = A_a^+ \exp\left(-\tilde{k}_a(x - t_c - t_w)\right) + \\ + A_a^- \exp\left(\tilde{k}_a(x - t_c - t_w)\right).$$

Here, we introduced the following notations:

$$\tilde{k}_c = \sqrt{2m_e(V_c - E)},$$
$$\tilde{k}_a = \sqrt{2m_e(V_a - E)},$$
$$k_0 = \sqrt{2m_e E}.$$

The wave function at the cathode represents an outgoing wave:

$$\psi(x) = A_c \exp(-ik_0 x).$$

Similarly, at the anode:

$$\psi(x) = A_a \exp\left(ik_0(x - d)\right).$$

Here, $A_a = k_0$, $d = t_c + t_w + t_a$ is the size of the structure. The task is to match the wave functions and their derivatives at the boundaries. There are eight unknowns, four boundaries, and thus eight conditions.

Calculating the eighth-order determinant without numerical methods is challenging, so we iteratively eliminate unknowns. The results of this elimination are provided in the Appendix. By defining the function $f(E)$ according to formula (A2) from the Appendix, the characteristic equation takes the form:

$$E = V_a \frac{\left(f\left(E\right)-1\right)^2}{\left(f\left(E\right)-1\right)^2 - \left(f\left(E\right)+1\right)^2}. \qquad (6)$$

This equation allows for the iterative search for complex roots $\tilde{E}_n$. Assuming the function f is large in magnitude (corresponding to wide barriers), we obtain $E \approx -V_c f\left(E\right)/4$. As the well expands from a very narrow width, the energy level first appears near V_a [21]. For such a level, the decay rate $\tilde{k}_a \approx 0$ is:

$$\mathrm{tg}\left(k_0 t_w\right) \approx k_0/\tilde{k}_a.$$

Assume there is such a level:

$$\tilde{E}_1 = V_a\left(1-\delta_1\right),$$

where δ_1 is small. Also, let:

$$\delta << \Delta = V_c/V_a - 1.$$

Calculating the function f, we obtain:

$$\tilde{k}_c \approx \tilde{k}\left(1+\delta/\left(2\Delta\right)\right),$$

$$\exp\left(2\tilde{k}_c t_c\right) \approx \exp\left(2\tilde{k}t_c\right)\left(1+\delta/\left(2\Delta\right)\right),$$

Where

$$\tilde{k} = \sqrt{2m_e V_a \Delta}/\hbar.$$

As well as:

$$k_0 \approx \tilde{k}\left(1-\delta/2\right)/\sqrt{\Delta},$$

$$k_0/\tilde{k}_c \approx \left(1-\delta_1\left(1+1/\Delta\right)/2+\delta_1^2/\left(4\Delta\right)\right)/\sqrt{\Delta}.$$

Let us set $\tilde{k}_a = 0$, and rewrite the introduced condition as

$$\left(t_w\tilde{k}/\sqrt{\Delta}\right) = \sqrt{\Delta}.$$

Then we have

$$f \approx \frac{\exp\left(2\tilde{k}t_c\right)}{4\Delta - \delta_1(\Delta+1)\left(1+t_w\tilde{k}\right)} \times$$

$$\times\{\delta_1(\Delta+1)(1+t_w\tilde{k}) -$$

$$-\left(\delta_1^2/2\right)\left[\Delta + 2\Delta\left(t_w\tilde{k}\right)^2 + (\Delta+1)\left(1+t_w\tilde{k}\right)\right]\}.$$

For the left side of equation (A2), we get

$$\frac{1+ik_0/\tilde{k}_a}{1-ik_0/\tilde{k}_a} \approx 1+\delta_1\frac{1+i\sqrt{\Delta}}{\left(\Delta+1\right)^2},$$

from which the correction δ_1 can be found, expanding further:

$$\delta_1 = \frac{16\exp\left(-2\tilde{k}t_c\right)}{\left(1+1/\Delta\right)\left(1+t_w\tilde{k}\right)} +$$

$$+\frac{4\delta\exp\left(-2\tilde{k}t_c\right)}{\left(1+1/\Delta\right)\left(1+t_w\tilde{k}\right)} \times$$

$$\times\left[4\frac{1+i\sqrt{\Delta}}{\left(\Delta+1\right)^2} - \left(1+1/\Delta\right)\left(1+t_w\tilde{k}\right)\right].$$

We can neglect the second-order term. To find the exact roots of equation (6), let us consider a well surrounded by infinitely wide barriers, i.e., potential steps of height V_c and V_a. In such a well, stationary energy levels $E_n < V_a$ are possible. The problem of an asymmetric well has been solved and studied in [21]. With the notation

$$k_{0n} = \sqrt{2m_e E_n}/\hbar$$

it has the solution

$$k_{0n}t_w = n\pi - \arcsin\left(\frac{k_{0n}\hbar}{\sqrt{2m_e V_c}}\right) -$$

$$-\arcsin\left(\frac{k_{0n}\hbar}{\sqrt{2m_e V_a}}\right) = g\left(E_n\right) \qquad (7)$$

which gives real energy levels. Rewrite equation (7) as

$$\mathrm{tg}\left(k_{0n}t_w\right) = k_{0n}t_n,$$

where

$$t_n = \frac{k_{0n}+\tilde{k}_c}{\tilde{k}_c\tilde{k}_a - k_{0n}^2}.$$

Choosing the well width from the condition of the existence of one level:

$$t_w > t_0 = \frac{\pi/2 - \arcsin\left(V_a/V_c\right)}{\sqrt{2m_e V_c}/\hbar},$$

we get

$$E_n = \frac{\left(\hbar g\left(E_n\right)/t_w\right)^2}{2m_e}.$$

From this equation, we find the real values of E_1 by the bisection method in the interval $\left[\left(0,V_0\right)\right.$. For the existence of multiple levels, the well must be several times wider than t_0. Then we find E_n, $n = 1,2,...,N$. The values of E_n are used as initial approximations $\tilde{E}_n^{(0)} = E_n$ for the iterations according to equation (6). As a result, we obtain all the levels from which a particle can escape to both the anode and the cathode. To increase the current, the widest possible well should be used, for which the electrode material should have the maximum electron mean free path (MFP). The MFP can be significantly increased by using cryogenic temperatures. Let us consider the derivation of equation (7), where the wave functions (WF) on the cathode and anode sides are taken as

$$\psi\left(x\right) = A_c \exp\left(\tilde{k}_a\left(1 - i\delta_c\right)x\right),$$

$$\psi\left(x\right) = A_a \exp\left(-\tilde{k}_a\left(1 - i\delta_a\right)\left(x - t_w\right)\right),$$

i.e., the barriers are partially transparent. Here

$$\tilde{k}_c = \sqrt{2m_e\left(V_c - E_n'\right)}/\hbar,$$

$$\tilde{k}_a = \sqrt{2m_e\left(V_a - E_n'\right)}/\hbar,$$

and small corrections are taken as

$$\delta_c = E_n''/\left(2V_c - 2E_n'\right),$$

$$\delta_a = E_n''/\left(2V_a - 2E_n'\right).$$

In reality, they are associated with the finite width of the barriers and the finite lifetime of the levels. In the well $0 < x < t_w$, we take

$$\psi\left(x\right) = A\sin\left(k_{0n}\left(1 + \delta_n\right)x + \delta\right),$$

where

$$\tilde{E}_n = \left(1 + \delta_n\right)^2\left(k_{0n}\hbar\right)^2/\left(2m_e\right),$$

and the small correction δ_n needs to be found. As a result, we obtain the characteristic equation for it:

$$\mathrm{tg}\left(k_{0n}\left(1 + \delta_n\right)t_w\right) = k_{0n}\left(1 + \delta_n\right)\times$$

$$\times\frac{\tilde{k}_a\left(1 - i\delta_a\right) + \tilde{k}_c\left(1 - i\delta_c\right)}{\tilde{k}_c\tilde{k}_a\left(1 - i\delta_a\right)\left(1 - i\delta_c\right) - k_{0n}^2\left(1 + \delta_n\right)^2}.$$

Introducing the notations

$$\delta'_{an} = \delta_a/\delta_n, \quad \delta'_{cn} = \delta_c/\delta_n.$$

Primed quantities are not small. Considering (8), to obtain the correction, expansion up to the second order in δ_n^2 should be used. We obtain $\delta_n = A_n/B_n$, where

$$A_n = \frac{t_w}{t_n} + k_{0n}^2 t_n t_w - 1 + \frac{i\left(\delta'_{an}\tilde{k}_a + \delta'_{cn}\tilde{k}_c\right)}{\tilde{k}_a + \tilde{k}_c} -$$

$$-\frac{2i\tilde{k}_c\tilde{k}_a\left(\delta'_{an} + \delta'_{cn}\right) + 2k_{0n}}{\tilde{k}_c\tilde{k}_a - k_{0n}^2},$$

$$B_n = \frac{\tilde{k}_c\tilde{k}_a\delta'_{an}\delta'_{cn} + k_{0n}^2}{\tilde{k}_c\tilde{k}_a - k_{0n}^2} - k_{0n}^2 t_n t_w +$$

$$+4\left[\frac{i\tilde{k}_c\tilde{k}_a\left(\delta'_{an} + \delta'_{cn}\right) + k_{0n}^2}{\tilde{k}_c\tilde{k}_a - k_{0n}^2}\right]^2 +$$

$$+\frac{2i\tilde{k}_c\tilde{k}_a\left(\delta'_{an} + \delta'_{cn}\right) + 2k_{0n}^2}{\tilde{k}_c\tilde{k}_a - k_{0n}^2}.$$

For the calculation of the correction, one can assume $\tilde{E}'_n = E_n$, and then

$$\tilde{E}_n''/E_n = -\Im\left(\delta_n\right),$$

while the real part also changes:

$$E_n' = E_n\left(1 + \mathrm{Re}\left(\delta_n\right)\right).$$

In Fig. 2, two levels are shown. From the cathode, tunneling to both levels with exit to the anode is possible. In this case, the cathode heats up because its level is above the Fermi level (Nottingham effect). When transitioning from the first level to the anode, the anode cools, while transitioning from the second level heats it up. Tunneling from the anode to the second level at T=0 is impossible. The lifetime of the level exponentially decreases with the narrowing of the barriers. The barriers narrow as the field U_g increases (Schottky effect), i.e., with an increase in well depth. At $U_g > U_a + E_{Fc}/\hbar$, stationary levels are possible in the well. Narrowing of the barriers also occurs with increasing voltage U and decreasing sizes t_c and t_a. There is a critical voltage at which the barrier relative to the Fermi level disappears, becoming nearly triangular. Indeed, using equation (1), where we denote

$$W' = W_c\frac{\left(1 - \alpha/d\right)\left(1 + \delta/d\right)^2}{\left(1 - \delta/d\right)^2\varepsilon},$$

assuming $d = t_g$ and neglecting small terms, this condition can be written as:

$$V(x_0) = E_{Fc} =$$

$$= E_{Fc} + W_c'\left(1 - \delta d/\left(x_0(d - x_0)\right)\right) - eU_g x_0/d.$$

From this, we find the point x_0 where this occurs. It is very close to the cathode, so we simplify the cubic equation by replacing $d - x_0$ with d:

$$x_0 = eU_g x_0^2/(W_c' d) + \delta.$$

Solving this quadratic equation iteratively, first assuming

$$x_0 = \delta$$

and then refining:

$$x_0 = \delta + \delta^2 eU_g/(W_c' d).$$

The refinement is very small, so we obtain the critical voltage:

$$U_g = W'(d/\delta - 1)/(2e) \approx W'd/(2e\delta).$$

For a work function of about 4 eV at d=2 nm, this corresponds to a critical electric field strength at the cathode of $2.35 \cdot 10^{10}$ V/m. Thus, in RT structures with well widths of a few t_0 and narrow barrier widths t_c and t_a, a significant increase in emission current is possible simply by increasing the size t_w. However, tunneling is ballistic transport without energy loss, so the width t_w must be significantly less than the electron mean free path (MFP) in the corresponding material. The characteristic size t_w at room temperature is a few nanometers. To reduce the lifetime of levels and increase current, the barriers should be made narrow. Their narrowing is also achieved by increasing electrode voltages. It is not difficult to obtain exact solutions to equation (6), but these equations are model-based. For real potentials (Fig. 1), one should solve the exact equations (5) or $R^{\pm}(E) = 0$. The table above presents the results of iterative calculations of complex energies. Calculations based on equation (5) and the conditions $R^{\pm}(E) = 0$ agree well.

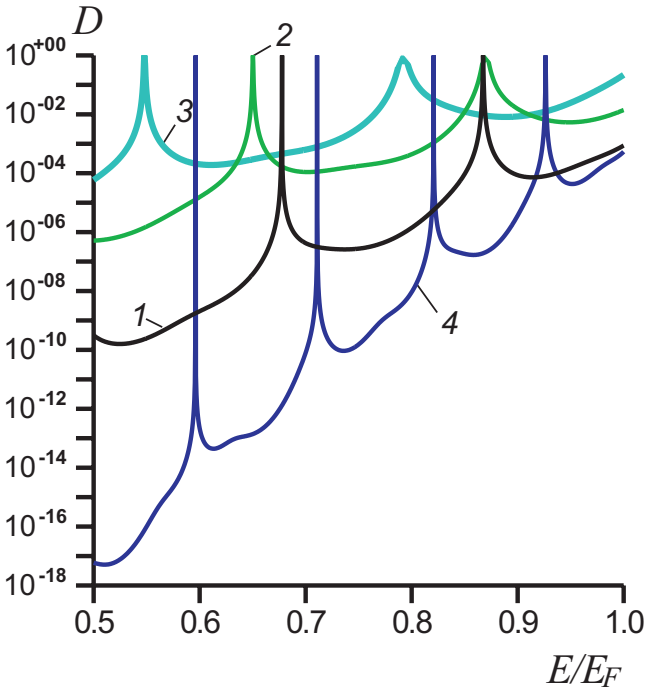

Fig. 3. Tunneling coefficient $D = D^+$ in a double-well RT structure as a function of the ratio $t = t_1 = t_2 = t_3$ depending on E/E_{F_c} at $t = t_g = 1$ nm (curves *1*, *3*) and $t = 2$ nm, $t_g = 1.5$ nm, $d = 9$ nm (curve *4*). Work function $W_c = W_g = W_a$ =4.0 eV, Fermi energy $E_{F_c} = E_{F_a}$ 5 eV, $U_a = 11$ V. $U_g = 13$ V (*1*, *4*); $U_g = 20$ V (*2*); $U_g = 25$ V (*3*)

A very simple method for determining complex levels is calculating the transparency of the structures. Figure 3 shows an example of calculating D^+ for several double-well RT structures with 2 to 4 metastable levels. Such structures are obtained with a double grid [1–3] and are more convenient for achieving resonance tunneling because two approximately equal barriers can be formed under a significant electrostatic potential U_a.

Table. Metastable levels (eV) in the range (0, E_{Fc}) for the potential in Fig. 1 at different anode voltages U_a (V):

U_a	1.0	2.0	3.0	4.0
$E_1' - iE_1''$	$0.14467 - i3.1 \cdot 10^{-4}$	$0.1445 - i2.9 \cdot 10^{-4}$	$0.1399 - i2.7 \cdot 10^{-4}$	$0.1405 - i2.7 \cdot 10^{-4}$
$E_2' - iE_2''$	$1.815 - i2.5 \cdot 10^{-3}$	$1.807 - i2.6 \cdot 10^{-3}$	$1.798 - i2.8 \cdot 10^{-3}$	$1.789 - i2.9 \cdot 10^{-3}$
$E_3' - iE_3''$	$4.4938 - i8.9 \cdot 10^{-3}$	$4.369 - i9.5 \cdot 10^{-3}$	$4.328 - i9.9 \cdot 10^{-3}$	$4.279 - i1.2 \cdot 10^{-2}$
$E_4' - iE_4''$		$6.872 - i7.2 \cdot 10^{-2}$	$6.982 - i8.3 \cdot 10^{-2}$	

Notably, the peaks for D^+ and D^- differ slightly, particularly at low energies. At $E > E_{Fc}$, $D^{\pm} \approx 1$ always holds. This case corresponds to thermionic emission if the electrode temperatures $T^{\pm} > 0$. It should be noted that for different barriers, peaks may not reach unity (incomplete resonance tunneling), associated with partial suppression of reflected electron waves. The values E_n' are determined by the peak maxima, which can be done accurately. The lifetimes E_n'' are determined by the resonance widths. Typically, the levels are located near the upper regions of the well.

Let us consider how the position and width of the level affect the current contribution. Suppose there is one level $E_1' - iE_1''$. Approximating it as an equilateral triangle with unit height, the contribution from the level is

$$\Delta J^+ = -em_e\left(E_{Fc} - E_1'\right)E_1''/\left(4\pi^2\hbar^3\right).$$

For levels near the Fermi level of the cathode, it is small. Therefore, it is important to obtain low-lying levels with a short lifetime (large width). For a single triangular potential barrier at a critical field, the semiclassical approximation gives its transparency D as

$$D \approx \exp\left(-4d\sqrt{2m_e}W^{3/2}/3\hbar eU_a\right),$$

see [20]. Here, the barrier height W is measured from the kinetic energy of the incoming electron, i.e., in our case, $W = V - E$.

For deep levels, the transparency of a single barrier is exponentially small compared to $D = 1$ in resonance tunneling. The formula works well for deep levels, but for $E = V$, its limitation becomes apparent: $D = 1$ at $W = 0$, while solving the Schrödinger equation gives $D < 1$. This limitation restricts the applicability of the Fowler–Nordheim formula to single barriers.

Nevertheless, the result can be used to estimate the lifetimes of deep levels by calculating $D_{(c,a)}$ at $W = V_{c,a} - E_n$ and determining $\delta_c = D_c$ and $\delta_a = D_a$.

3. LIFETIME OF THE LEVEL IN THE NONSTATIONARY APPROACH

The nonstationary Schrödinger equation (SE) is written as

$$\hat{S}(t,x)\psi(t,x) = V(t,x)\psi(t,x),$$

It is known to be relativistically non-covariant. Here, the operator for a free particle's Hamiltonian is denoted as

$$\hat{S}(t,x) = i\hbar\partial_t + \frac{\left(\hbar\partial_x\right)^2}{2m_e}.$$

This implies that the Green's propagator function (GPF), which describes the propagation of a particle from point x' at time t' to point x at time t, has the following form [22, 23]:

$$K_0(t-t',x-x') = \\ = \operatorname{sgn}(t-t')\sqrt{\frac{m_e}{2\pi i\hbar|t-t'|}} \times \\ \times\exp\left(\frac{i(x-x')^2 m_e}{2\hbar|t-t'|}\right), \qquad (8)$$

This expression suggests infinitely fast propagation of the perturbation. Indeed, GPF (8) defines the particle's presence at point x at time t based on its amplitude $\psi_0(x',t')$ at point x' at the initial moment t':

$$\psi_0(t,x) = \int_{-\infty}^{\infty} K_0(x-x',t-t')\psi_0(t',x')dx'.$$

If at the initial moment t_0, a probability density

$$\psi_0(t_0,x,) = \delta(x-x_0),$$

emerges at point x_0 meaning the particle is localized there, then for any later time $t > t_0$, the wavefunction exists throughout the entire infinite space:

$$\psi_0(x,t) = K_0(x-x_0,t-t_0),$$

Thus, the propagation speed of the probability density is infinite, though the density itself rapidly decreases at distant points. Here, the subscript "0" denotes a free particle (V=0).

Such a particle is generally described as a wave packet (WP) with a certain spectrum of wave numbers k and energies E. It is worth noting that the incoming particle flow described by distribution (2) also represents a WP.

The GPF (8) satisfies the initial condition

$$K_0(t-t',x-x')_{t\to t'} = \delta(x-x')$$

and the differential equation

$$\hat{S}K_0(t-t',x-x') = i\hbar\delta(t-t')\delta(x-x'),$$

(see [22]). Solving the nonstationary SE requires setting appropriate initial conditions. A convenient approach is to use the stationary case at the initial moment $t = t_0$, i.e., the wavefunction $\psi_0(x)$ and the potential $V_0(x)$.

At $t \geq t_0$, when the potential $V(x,t)$ starts changing, the wavefunction satisfies the nonstationary SE. The wavefunction for $t_0 > 0$ is governed by the Lippmann–Schwinger-type integral equation:

$$\psi(x,t) = \psi_0(x) - i\hbar^{-1}\int_0^t\int_{-\infty}^{\infty} K_0(t-t',x-x')\times$$
$$\times\left[V(t',x') - V_0(x')\right]\psi(t',x')dx'dt'. \quad (9)$$

Indeed, at t " 0,, we have $\psi(x,t) = \psi_0(x)$.. Taking the time $t > 0$ derivative of (9) and applying the operator $\hat{S}$, we obtain the SE:

$$\hat{S}\psi(x,t) = V(t,x)\psi(t',x').$$

Assume the potential

$$\Delta V(t,x) = V(t,x) - V_0(x)$$

is localized within a certain region. In this case, for small times, equation (9) can be solved rather simply. An example for a double-barrier resonant tunneling diode (RTD) is provided in [12]. This equation is particularly convenient for analyzing transient processes and tunneling times. Two cases can be considered: (a) $\Delta V(0,x) = 0$ (smooth potential change) and (b) $\Delta V(0,x) \neq 0$. We focus on the second case here. Assume a well with one metastable level between two barriers exists at $t < 0$. This level cannot be populated, as it would decay over infinite time. For simplicity, consider identical barriers of height V. The metastable level between identical barriers V is defined by the condition

$$\mathrm{th}\left(\tilde{k}_1 t_b\right)(kt_w) =$$
$$= \alpha = \frac{\sqrt{\tilde{E}_1\left(V-\tilde{E}_1\right)}}{\tilde{E}_1 - V/2},$$

see [13], where

$$\tilde{k}_1 = \tilde{k}'_1 + i\tilde{k}''_1 = \frac{\sqrt{2m_e\left(V-\tilde{E}_1\right)}}{\hbar}$$

This equation determines the level's lifetime, $t_1 = t_2 = t_b$ is the barrier width. A convenient numerical solution can be sought in the form

$$E_1 = V_0{}^2\left(\alpha(E_1)\right)/\mathrm{th}^2\left(\tilde{k}_1 t_b\right),$$

$$V_0 = \hbar^2/2m_e t_w^2,$$

expressing the arctangent via logarithmic functions. The quantities

$$\alpha = \alpha' + i\alpha'' = \sqrt{E_1(V-E_1)}/(E_1 - V/2)$$

and

$$\tilde{k}' = \tilde{k}'_1 + i\tilde{k}''_1$$

are complex. For wide barriers, we obtain

$$\tilde{k}'_1 = \sqrt{2m_e(V-E'_1)}/\hbar,$$

$$\tilde{k}''_1 = E''_1\sqrt{m_e/(2V-2E'_1)}/\hbar,$$

$$\mathrm{th}^2\left(\tilde{k}_1 t_b\right) \approx 1 - 4\exp\left(-2\tilde{k}'_1 t_b\right)\exp\left(-2i\tilde{k}''_1 t_b\right).$$

It is easiest to estimate the level by assuming it arises at the barrier boundary. In this case, α', and

$$\alpha'' \approx 3\alpha'_1 E''_1/2V$$

To simplify further calculations, introduce the dimensionless parameter

$$\delta = 4\exp\left(-2\tilde{k}'_1 t_b\right)\exp\left(-2i\tilde{k}''_1 t_b\right),$$

and obtain the energy as

$$E_1 = (V+V_0)/2 + \sqrt{V_0^2/4 + 3V_0V/2} + \delta\Delta E_1,$$

where

$$\Delta E_1 = \frac{\left(V_0{}^2/4 + 3VV_0/4\right)}{\sqrt{V_0^2 4 + 3V_0V/2}} + \frac{V_0}{2}.$$

If the well deepens by an amount ΔV, the energy at the bottom becomes negative. Assume only one stable level exists. If the center of the well is at x=0, the wavefunction inside the well takes the form of either an even or odd function:

For an even wavefunction:

$$\psi_0(x) = A_c\cos(k_1 x),$$

For an odd wavefunction:

$$\psi_0(x) = A_s \sin(k_1 x),$$

herewith

$$\psi_0(t_w/2) \neq 0, \quad \psi_0'(t_w/2) \neq 0,$$

$$k_1 = \sqrt{2m_e E_1}/\hbar.$$

Let's mark

$$\tilde{k}_0 = \sqrt{2m_e(\Delta V - E_1)}/\hbar,$$

$$\tilde{k}_1 = \sqrt{2m_e(V + \Delta V - E_1)}/\hbar,$$

$$z_0 = -i/\tilde{k}_0,\ \tilde{z}_1 = -i/\tilde{k}_1,\ z_1 = 1/k_1.$$

Then, in the case of an odd wave function, we have the characteristic equation:

$$\mathrm{tg}(k_1 t_w/2) = iZ_i/\rho_1,$$

And for the even function:

$$\mathrm{tg}(k_1 t_w/2) = -i\rho_1/Z_i.$$

The value

$$Z_i = \frac{\tilde{\rho}_1\left(\tilde{\rho}_0 - \tilde{\rho}_1 \mathrm{th}\left(\tilde{k}_1 t\right)\right)}{\tilde{\rho}_1 - \tilde{\rho}_0 \mathrm{th}\left(\tilde{k}_1 t\right)}$$

is imaginary, so the equations are real and determine the real energies. We take the normalization of the wave function (WF) from the condition of finding the particle in the well region $|x| \leq t_w/2$. This is an approximate condition, as there is some probability leakage through the barriers. However, with sufficiently wide barriers, it is negligible. A strict normalization can be performed, but it results in cumbersome amplitude values. In our case, the amplitudes are:

$$|A_c|^2 = \left[t_w\left(1 + \sin(t_w k_1)/(t_w k_1)\right)\right]^{-1},$$

$$|A_s|^2 = \left[t_w\left(1 - \sin(t_w k_1)/(t_w k_1)\right)\right]^{-1}.$$

It is clear that the even level should appear first, as its wave function approximately corresponds to the half-wave of de Broglie. Thus, for $t < 0$, such a populated level exists. At the moment $t = 0$, the potential $\Delta V > 0$ is suddenly switched on, and the bottom of the well rises to zero energy. In such a well, the particle cannot exist indefinitely, and the state begins to decay, described by the integral equation (IE):

$$\psi(x,t) = \psi_0(x) - i\hbar^{-1}\Delta V \times$$

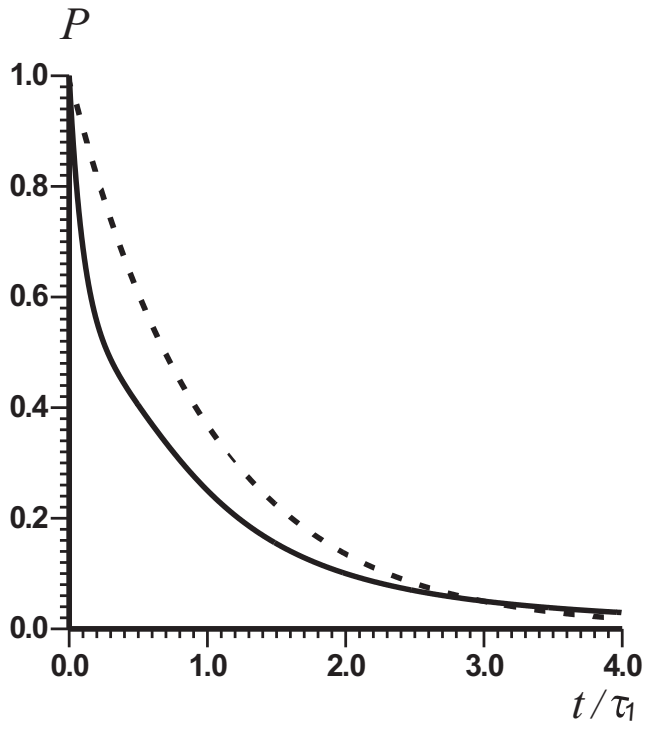

Fig. 4. Transition probability $P(t)$ according to formula (12) for the decay of a single level. The dashed curve represents exponential decay $P_0(t) = \exp(-t/\tau_1)$.

$$\times \int_0^t \int_{-t_w/2}^{t_w/2} K_0(t - t', x - x')\psi(t', x')\,dx'dt'. \qquad (10)$$

This problem can be solved numerically or by perturbation theory. In the latter case, the first approximation is:

$$\psi_{(1)}(x,t) = \psi_0(x) - i\hbar^{-1}\Delta V \times$$

$$\times \int_0^t \int_{-t_w/2}^{t_w/2} K_0(t - t', x - x')\psi_0(x')\,dx'dt'.$$

The probability of finding the particle in the well region now becomes:

$$P(t) = \int_{-t_w/2}^{t_w/2} |\psi(t,x)|^2\,dx. \qquad (11)$$

It decreases over time. Solving IE (10), we compute (11). Obviously, with the chosen normalization $P(0) = 1$. Approximating (11) with the function $P_0(t) = \exp(-t/\tau_1)$, we determine the level lifetime. The corresponding result is shown in Fig. 4, corresponding to the value $E_1''/E_1' = 0.021$,

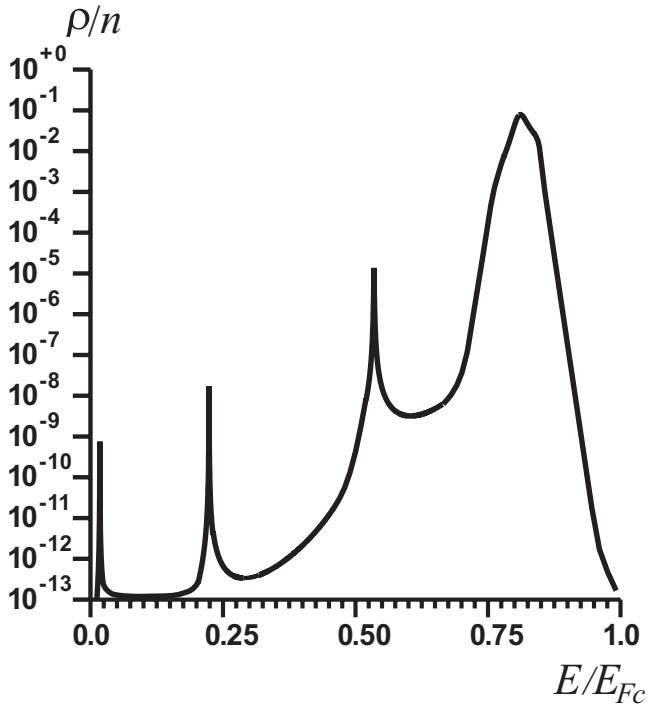

Fig. 5. Normalized particle number density ρ in the well as a function of energy E for three resonance levels (eV): 0.140552, 1.78936 eV, 4.27933 eV (see Table, $U_a = 4$ V)

$\tau_1 = 59.5$ fs. It should be noted that the decay of such a state generally does not follow an exponential law [24–35], which holds only for infinitely long-lived levels [36]. There exist continuous-spectrum states in the well that distort the exponential law. The continuous spectrum and interference lead to faster initial decay, followed by a slowdown [17, 24]. An even more complex case corresponds to multiple levels. The non-stationary approach is significantly more complicated than determining complex roots. Interestingly, for tunneling problems, calculating the probability density:

$$\rho(E) = \int \left|\psi_0(x,E)\right|^2 dx$$

both in the well region and in the barrier region shows maxima at energies corresponding to the resonance levels E′ (see Fig. 5). The result is normalized to the particle number density in both flows:

$$n(E) = n^+(E) + n^-(E) =$$
$$= \sqrt{2E} m_e^{3/2} (E_{Fc} - E)/\pi^2\hbar^3.$$

This is because all incident flows from the left and right with resonance energies E_n' pass into the well, while for other energies they are significantly reflected. Both the tunnel current density J and the probability current density j are continuous along the entire structure, including the electrodes, reflecting the conservation law of particles (probability) in non-relativistic quantum mechanics.

4. APPLICATION OF NON-STATIONARY SCHRÖDINGER EQUATION FOR DETERMINING TUNNELING TIME

Since the introduction of the concept of tunneling time in 1930, there has been no established understanding in the literature (see [19] and references therein). Paradoxes such as the Hartman effect, "superluminal" tunneling, negative tunneling time, and others are still discussed. IE (10) is quite convenient for resolving such issues and studying transient processes [12]. The level lifetime (residence time in the structure) is often associated with tunneling time. Here, instead of IE (9), we consider another approach based on series expansion for solving the non-stationary Schrödinger equation. Suppose that at $t < 0$, in the region $0 < x < d$, we have a structure with three electrodes: $U_a = 0$, and $U_g = -W_c/e$. Also, let the value ddd be sufficiently large. In this case, the potential is close to a rectangular shape with width d and height W_c relative to the Fermi level (see Figure 1, curve 0). Relative to zero, its height is $V = W_c + E_{Fc}$. If the grid voltage were zero, the potential (relative to E_{Fc}) would appear as two peaks of height W_c, separated by a gap with zero height. Curve 1 in Fig. 1 demonstrates the potential at a negative grid voltage $U_g = -W_c$, when the entire curve is elevated by W_c. Such a potential blocks the current. Suppose that at time $t = 0$, the potentials switch such that $U_a > 0$ and $U_g = E_{Fc}/e$, i.e., the problem becomes equivalent to resonant tunneling (RT). Accordingly, we need to consider the transient processes of tunneling establishment when $t > 0$ during the switch from curve 1 to curves like 2, 4. For a diode structure, this switch corresponds to curve 0 transitioning to curves 3, 5, 7, but without RT. The macroscopic change in current during such a process is quite easy to measure, unlike the tunneling time of an individual particle. It should be noted that for $t < 0$, the current was absent due to the symmetry of the structure. Also, at these times, the particle density in the structure was negligible, as the tunneling probability through a wide barrier was nearly zero. Near the edges, the density decays exponentially. By choosing a large ddd, one can assume that particles were absent in the barrier region. Switching the potentials leads to the appearance of current. It cannot

appear instantaneously, as particles need to traverse the region d, thus creating a finite transient time. We will solve the non-stationary Schrödinger equation by expanding into series in the region $0 < x < d$:

$$\Psi(t,x) = $$

$$= A\left(\sum_{n=0}^{\infty}\alpha_n(t)\cos(\chi_n x)+\sum_{n=1}^{\infty}\beta_n(t)\sin(\chi_n x)\right), \quad (12)$$

$$V(t,x) = \sum_{n=0}^{\infty}\nu_n(t)\cos(\chi_n x), \quad (13)$$

where $\chi_n = n\pi/d$. This method is applicable for multiple electrodes, but further numerical results are presented for the diode. It is not possible to use only cosines or sines in the expansion (12), as this would always result in zero probability current density. For simplicity, we will apply the method to the diode structure. The amplitude A is introduced for normalization, meaning that when it is specified, we can assume $\alpha_0 = 1$. To perform the calculations, we truncate the series (12) and (13) by an index N. Substituting (12) and (13) into the Schrödinger equation and using the orthogonality of trigonometric functions, we obtain the coupled differential equations:

$$\alpha_n'(t) = i\sum_{m=0}^{N}A_{nm}^{cc}\omega_m\alpha_m(t) +$$

$$+ i\sum_{m=1}^{\infty}A_{nm}^{cs}\omega_m\beta_m(t)\sin(\chi_m x) -$$

$$-\frac{i}{\hbar}\left(\sum_{m=0}^{N}V_{nm}^{\alpha cc}(t)\alpha_m(t) +\right.$$

$$\left.+\sum_{m=1}^{\infty}V_{nm}^{\beta cs}(t)\beta_m(t)\sin(\chi_m x)\right), \quad (14)$$

$$\beta_n'(t) = i\sum_{m=0}^{N}A_{nm}^{sc}\omega_m\alpha_m(t) +$$

$$+i\sum_{m=1}^{\infty}A_{nm}^{ss}\omega_m\beta_m(t)\sin(\chi_m x) -$$

$$-\frac{i}{\hbar}\left(\sum_{m=0}^{N}V_{nm}^{\alpha sc}(t)\alpha_m(t) +\right.$$

$$\left.+\sum_{m=1}^{\infty}V_{nm}^{\beta ss}(t)\beta_m(t)\sin(\chi_m x)\right). \quad (15)$$

Here, $\omega_n = \hbar\pi^2 n^2/\left(2m_e d^2\right)$ are the frequencies, and the matrix elements, detailed in the Appendix. These equations are quite complex if the potential depends arbitrarily on time. In the case of an abrupt potential switch, it stops depending on time, simplifying the equations. Rewriting the matrix elements, the first equation can be simplified to:

$$\alpha_n'(t) - i\omega_n\alpha_n(t) =$$

$$= f_n(t) = -\frac{i}{\hbar}V_{nn}^{\alpha cc}(t) + i\sum_{m=1}^{\infty}\tilde{A}_{nm}^{cs}(t)\beta_m(t) +$$

$$+i\sum_{m=0,m\approx n}^{N}\tilde{A}_{nm}^{cc}(t)\alpha_m(t).$$

Solving this equation using the Bernoulli method or the method of variation of arbitrary constants gives:

$$\alpha_n(t) = \alpha_n(0)\exp(i\omega_n t) +$$

$$+\exp(i\omega_n t)\int_0^t f_n(t')\exp(-i\omega_n t')dt'. \quad (16)$$

Similarly, we obtain:

$$\beta_n(t) = \beta_n(0)\exp(i\omega_n t) +$$

$$+\exp(i\omega_n t)\int_0^t g_n(t')\exp(-i\omega_n t')dt'.$$

Here, the following functions are introduced:

$$g_n(t) = -\frac{i}{\hbar}V_{nn}^{\beta ss}(t) + i\sum_{m=0}^{N}\tilde{A}_{nm}^{sc}(t)\alpha_m(t) +$$

$$+i\sum_{m=1,m\approx n}^{\infty}\tilde{A}_{nm}^{ss}(t)\beta_m(t).$$

The solution in time is sought using the discretization method: $t_m = m\Delta t$, $m = 1,2,...$, with integrals calculated using the midpoint method. If the initial values $\alpha_n(0)$, $\beta_n(0)$, are known, the equations allow us to find $\alpha_n(m\Delta t)$, $\beta_n(m\Delta t)$, using either explicit or implicit schemes.

The modified matrix elements here take a simple and clear form, for example:

$$\tilde{A}_{nm}^{\alpha cc}(t) = A_{nm}^{cc} - V_{nm}^{\alpha cc}(t)$$

If such a barrier instantaneously changes its shape at $t = 0$ to $V(x)$, these elements stop depending on time:

$$\tilde{A}_{nm}^{\alpha cc} = \omega_m A_{nm}^{cc} - V_{nm}^{\alpha cc}/\hbar$$

Their exact values can be found if the shape $V(x)$ is simple. For large U_a, it resembles a triangle placed on a rectangular base (see Fig. 1, curves 3, 5, 7). At $eU_a = E_F$, the height of the base can be taken as W, and the height of the triangle as E_F. Due to the Schottky effect, the barrier is actually somewhat lower. Calculating the integrals, we obtain:

$$V_{nm}^{\alpha cc} = \left(v_{n+m}^{c} + v_{n-m}^{c}\right)/\left(1+\delta_{n0}\right),$$

$$V_{nm}^{\beta ss} = \left(v_{n+m}^{s} - v_{n-n}^{s}\right)/\left(1+\delta_{n0}\right),$$

$$V_{nm}^{\alpha sc} = v_{n-m}^{s} + v_{n+m}^{s},$$

$$V_{nm}^{\beta ss} = v_{m-n}^{c} - v_{m+n}^{c}.$$

Here, the values of the following integrals are introduced:

$$\nu_n^c = \frac{1}{d}\int_0^d V(x)\cos\left(\chi_n x\right)dx,$$

$$\nu_n^s = \frac{1}{d}\int_0^d V(x)\sin\left(\chi_n x\right)dx.$$

For the initial symmetric wide barrier (curve 0), the height $V_0 = W + E_F$, and the coefficients $\alpha_n(t) = 0$ and $\beta_n(t) = 0$ at $t < 0$, as the probability density inside is practically absent. This approximation improves with increasing d, implying $\alpha_n(t) = 0$, $\beta_n(t) = 0$, i.e. within $\Psi(x,t) = 0$, $t < 0$. We take the initial barrier as rectangular. Then the integrals are easily computed, for example:

$$\nu_n(0) = V_0 sinc(n\pi) = V_0\delta_{n0}.$$

When this barrier under applied voltage $U_a = E_F/e$ takes the form:

$$V(x) \approx W + E_F\left(1 - x/d\right)$$

(see Fig. 1, curve 7), we obtain:

$$\nu_n^c = \left(W + E_F\right)\delta_{n0} + E_F conc(n\pi),$$

$$\nu_n^s = W conc(n\pi) + E_F sinc(n\pi).$$

In our case:

$$\nu_n^s = W conc(n\pi).$$

We assume that at the moment of voltage application, some coefficients $\alpha_n(0)$ and $\beta(0)$ instantly change from zero. This happens due to the appearance of probability current density. We find them from the continuity condition of this current density.

To the left of the barrier, the spectral wave function has the form:

$$\psi(x,k) = a^+(k)\left[\exp(ikx) + R^+(k)\exp(ikx)\right],$$

and to the right:

$$\psi\left(x,\tilde{k}\right) =$$

$$= a^-\left(\tilde{k}\right)\left[\exp\left(-i\tilde{k}\left(x-d\right)\right) - \exp\left(i\tilde{k}\left(x-d\right)\right)\right].$$

Here:

$$\psi(d,k) = 0, \quad \psi'\left(d,\tilde{k}\right) = -2i\tilde{k}a^-\left(\tilde{k}\right),$$

At high voltage:

$$\left|\psi'(d,k)/\psi'\left(0,\tilde{k}\right)\right| << 1.$$

Upon voltage application, the electrochemical potential on the cathode jumps, hence:

$$\sqrt{2m_e eU_a}/\hbar < k < \sqrt{2m_e\left(E_{Fc} + eU_a\right)}/\hbar,$$

$$0 < \tilde{k} < \sqrt{2m_e E_{Fa}}/\hbar.$$

Now the coefficients α_n, β_n in the wave function (13) at $t > 0$ become non-zero. They are dimensionless, so the amplitude A must be determined from the normalization to particle flux. The flux to the right, at large U_a, can be taken as zero:

$$j(d) = 0.$$

The flux to the left for the wave function:

$$\psi(x,k) = a^+(k)\left[\exp(ikx) + R^+(k)\exp(ikx)\right]$$

is given by:

$$j(0,k) = \frac{\hbar k\left|a^+(k)\right|^2}{m_e}\left(1 - \left|R^+(k)\right|^2\right).$$

The total flux is obtained by integration:

$$j(0) = \frac{\hbar}{m_e}\int_0^{k_F}\left|a^+(k)\right|^2\left(1 - \left|R^+(k)\right|^2\right)kdk =$$

$$= \frac{m_e}{2\pi^2\hbar^3}\int_0^{E_F}\left(1 - \left|R^+(E)\right|^2\right)\left(E_F - E\right)dE.$$

Calculating the flux into the barrier from the left at $x = 0$, we find the condition:

$$j(0) = -i\frac{\hbar}{m_e}\mathrm{Re}\,\Psi'(0,0)\Psi^*(0,0) =$$

$$= -\frac{\hbar|A|^2}{m_e}\Re\left(i\sum_{n=1}^{\infty}\chi_n\beta_n(0)\right)\left(\sum_{m=0}^{\infty}\alpha_m^*(0)\right).$$

For the flux on the right (from the anode), we find:

$$-\frac{\hbar|A|^2}{m_e}\mathrm{Re}\left(i\sum_{n=1}^{\infty}(-1)^n\chi_n\beta_n(0)\right)\times$$

$$\times\left(\sum_{m=0}^{\infty}(-1)^m\alpha_m^*(0)\right) = 0.$$

It is also necessary to equate the wave functions (WF) and their derivatives at the boundaries of the region:

$$\Psi(0,0) = A\sum_{n=0}^{\infty}\alpha_n(0),$$

$$\Psi'(0,0) = A\sum_{n=1}^{\infty}\chi_n\beta_n(0),$$

$$\Psi(0,d) = A\sum_{n=0}^{\infty}(-1)^n\alpha_n(0) = 0,$$

$$\Psi'(0,d) = A\sum_{n=1}^{\infty}(-1)^n\chi_n\beta_n(0) = 0.$$

The last equality is set to zero because we assume a high voltage and measure the energy from the conduction band edge of the cathode. We obtain six additional equations to determine the infinite number of initial conditions $\alpha_n(0)$, $\alpha_n(0)$ $\beta_n(0)$. However, using the full set of sines in (12) is redundant because the cosine system is sufficient for approximating the wave function. We introduced sines to obtain nonzero fluxes and nonzero WF derivatives at the boundaries. It is quite reasonable to assume: $\alpha_n(0) = 0$, $n > 2$, $\beta_n(0) = 0$, $n > 3$. Thus, we have six unknowns, as well as six conditions. It is sufficient to consider nonzero coefficients $\alpha_0(0)$, $\alpha_1(0)$, $\beta_1(0)$, $\beta_2(0)$. Then:

$$\alpha_1(0) = \alpha_0(0), \quad \beta_2(0) = \beta_1(0)/2,$$

and all six equations are satisfied, with:

$$j(0) = -\frac{4\pi\hbar|A|^2}{m_e d}\mathrm{Re}\left(i\beta_1(0)\alpha_0^*(0)\right).$$

It is convenient to choose:

$$\beta_1(0) = i\alpha_0(0), \quad \alpha_0(0) = 1.$$

Then:

$$j(0) = \frac{4\pi\hbar|A|^2}{m_e d},$$

and the wave function takes the form:

$$\Psi(t,x) = \sqrt{\frac{j(0)m_e d}{4\pi\hbar}}\times$$

$$\times\left(\sum_{n=0}^{\infty}\alpha_n(t)\cos(\chi_n x) + \sum_{n=1}^{\infty}\beta_n(t)\sin(\chi_n x)\right). \qquad (17)$$

From this, we find $\Psi(t,d)$ and $\Psi'(t,d)$.

Another method for solving equations (14) and (15) involves Fourier transforms:

$$\begin{pmatrix}\alpha_n(t)\\ \beta_n(t)\end{pmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}\alpha_n(\omega)\\ \beta_n(\omega)\end{pmatrix}\exp(i\omega)d\omega,$$

which requires calculating integrals. This can be done using the residue method, but this approach requires separate consideration. To solve the problem, we need to determine the initial wave function $\Psi(0,x)$ and its derivative, which will be done below. It is convenient to introduce the frequency $\omega = E/\hbar$. The incident wave packet (WP) from the left can be written as:

$$\Psi(t,0) = \int_0^{E/\hbar}\psi^+(0,\omega)\exp(-i\omega t)d\omega,$$

$$\psi^+(0,\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Psi(0,\omega)\exp(i\omega t)d\omega.$$

Here:

$$\omega = k^2\hbar/2m_e, \quad k = \sqrt{2m_e\omega/\hbar}.$$

Neglecting back tunneling, we have on the left:

$$\Psi(t,0) = \int_0^{E/\hbar}a^+(\omega)\left(1 + R^+(\omega)\right)\exp(-i\omega t)d\omega,$$

and on the right:

$$\Psi(t,d) = \int_0^{E_F/\hbar}a^+(\omega)T^+(\omega)\exp(-i\omega t)d\omega.$$

The incident WP from the left is denoted as:

$$\Psi^+(t) = \int_0^{E/\hbar}a^+(\omega)\exp(-i\omega t)d\omega.$$

Here:

$$\psi^+(0,\omega) = a^+(\omega), \quad \psi^+(d,\omega) = a^+(\omega)T^+(\omega).$$

Defining $\Psi(x,t)$ as the solution to the nonstationary Schrödinger equation at time t, we construct the function:

$$\tilde{\Psi}(x,t) = \Psi(x,t) - \Psi(x,0).$$

It is zero outside the interval $(0,t)$, meaning it has a limited support, and:

$$\tilde{\Psi}(x,\omega) = \frac{1}{2\pi}\int_0^t \tilde{\Psi}(x,t')\exp(i\omega t')dt',$$

$$\Psi(x,\omega) = \frac{1}{2\pi}\int_{-\infty}^t \Psi(x,t')\exp(i\omega t')dt'.$$

We can construct the time-dependent reflection and transmission coefficients $R^+(t)$, $T^+(t)$. Specifically, we take:

$$R^+(t) = \Psi(0,t)/\Psi^+(t) - 1,$$

$$T^+(t) = \Psi(d,t)/\Psi^+(t).$$

Considering back tunneling, we define the incident WP from the right:

$$\Psi^-(t) = \int_0^{E/\hbar} a^-(\omega)\exp(-i\omega t)d\omega.$$

Thus, we obtain:

$$\Psi(t,0) = \Psi^+(t)\left(1 + R^+(t)\right) + T^-(t)\Psi^-(t),$$

$$\Psi(t,d) = \Psi^+(t)T^+(t) + \Psi^-(t)\left(1 + R^-(t)\right).$$

To find all coefficients, we also need to determine $\Psi'(t,x)$,Ψ'^+ and Ψ'^-. Derivatives can be found by differentiating the series. The current density at the anode is defined through the probability current density:

$$J(t_m) = -ej(t_m,d).$$

For this, when normalizing the wave function to the probability density, we use [20]:

$$j(t_m,x) =$$

$$= -\frac{i\hbar}{2m_e}[\Psi^*(t_m,x)\partial_x\Psi(t_m,x) -$$

$$-\Psi(t_m,x)\partial_x\Psi^*(t_m,x)].$$

For an arbitrary moment in time, we obtain:

$$j(t,x) = \frac{j(0)}{4}\times$$

$$\times\mathrm{Re}(-i\sum_{m=0}^{\infty}\left[\alpha_m^*(t)\cos(\chi_m x) + \beta_n^*(t)\sin(\chi_m x)\right]\times$$

$$\times\sum_{n=1}^{\infty} n\left[-\alpha_n(t)\sin(\chi_n x) + \beta_n(t)\cos(\chi_n x)\right]),$$

$$j(t,d) = \frac{j(0)}{4}\times$$

$$\times\mathrm{Re}\left(-i\sum_{m=0}^{\infty}(-1)^m\,\alpha_m^*(t)\cdot\sum_{n=1}^{\infty}(-1)^n\, n\beta_n(t)\right).$$

From this equation, it follows that:

$$j(0,d) = 0, \quad j(\Delta t,d) \sim \Delta t,$$

i.e., instantaneous tunneling and negative tunneling time are not possible. Using the spectra $\Psi(d,\omega)$ and $\Psi'(d,\omega)$, the result can be represented as:

$$j(t,d) = \frac{\hbar}{(2\pi)^2 m_e}\times$$

$$\times\,\mathrm{Re}\int_{-\infty}^{\infty}(-i)\Psi^*(d,\omega)\Psi'(d,\omega')\exp\left(i(\omega'-\omega)t\right)d\omega' d\omega.$$

For the steady-state process, the spectral wave function at the anode is:

$$\psi(x,k) = a^+(k)T^+(k)\exp\left(ik_a(x-d)\right)$$

The probability flux density for this wave function is:

$$dj(d,k) = v_a(k)\left|a^+(k)T^+(k)\right|^2 dk,$$

where the speed at the anode is:

$$v_a(k) = \sqrt{v^2(k) + 2eU_a/m_e}.$$

It should be noted that this speed is greater than $v(k)$ due to the acceleration of electrons passing through the barrier by the anode. Over the free path length, they scatter and transition to the Fermi level of the anode, with $v_a(k)$ decreasing to $v(k)$, causing the anode to heat up. The method of series used here is also convenient for solving the Schrödinger equation (SE) together with the Poisson equation (PE).

5. RESULTS AND DISCUSSION

Figs. 6 and 7 present the results of the transient process calculations, showing the establishment of the anode current in a diode with a Fermi energy of 7 eV and the probability density distribution $|\Psi(x,t)|^2$ when stepwise voltages of 3, 5, and 7 V appear at the anode. Fig. 7 shows the probability density distribution for curve 1 of Fig. 6 at different moments in time. The oscillations in probability density result from the finite sums used in the calculations. As the number of terms in the sums increases, both the oscillation amplitude and period decrease. The SE was integrated using the series method with 40 terms and an explicit calculation of the coefficients in equation (12). Expanding in other bases in (12) allows eliminating the oscillations. For example, finite elements can be used. However, the proposed series method is convenient when solving the SE and PE simultaneously, as applied in [2].

Calculations were performed using 200 time points. Curves 2 and 3 in Fig. 6 were constructed using 50 time points. For copper (Fermi energy 7 eV), we have an electron concentration of $8.5 \cdot 10^{28} \mathrm{m}^{-3}$ and a Fermi velocity $v_F = 1.57 \cdot 10^6$ m/s, meaning that a particle with this speed travels a distance $d = 10$ nm in a time $\tau = 6.35$ fs. We assumed that at the moment the voltage is applied, the probability density inside the barrier was zero. More precisely, it is symmetrically distributed relative to the center, approximately following a hyperbolic cosine distribution, increasing towards the edges, but extremely small at the edges themselves due to the near-complete reflection by the wide, nearly rectangular barrier. In this case, there is no inward probability flux into the barrier. The results shown in Fig. 6 indicate that the average transport speed of the probability density is somewhat greater than v_F, leading to the conclusion that the movement of the probability density is a collective effect caused by the interference of partial waves of the wave packet. An electron inside the barrier, or generally within a potential field, behaves as a quasiparticle defined by its interaction with many other particles. This averaged interaction determines the potential. A clear example is the potential of the image method. Such a quasiparticle is not required to behave like a free electron. Additionally, after passing the turning point for a single barrier, the electron moves quasi-classically and is accelerated by the anode. The additional velocity gained at $U_a = 5$ V is 1.33×106 m/s, approximately equal to

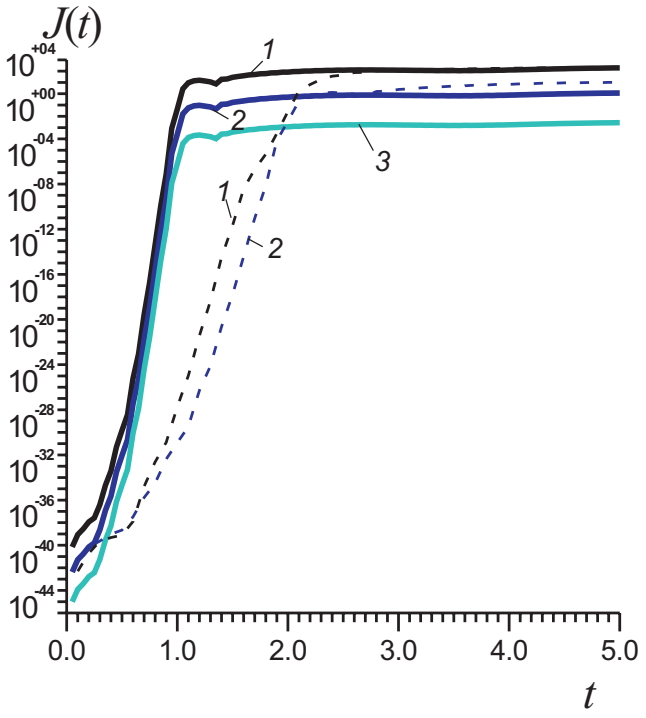

Fig. 6. Transient processes ($-J$ in A/cm², time in fs) during switching from the nearly rectangular barrier *1* to barriers *2*, *3*, and *4* in Fig. 1 (corresponding to curves *1*, *2*, and *3*).

v_F. Accordingly, the transit time is halved. A similar problem for resonant tunneling (RT) leads to a significantly longer transient process time. This can be explained by the need to form reflections from the barriers for RT to occur.

Formally, the lifetimes of the levels can be considered as an additional contribution to the transient process time. In Fig. 6, it is evident that the probability density is very small at short times. This function is asymmetric and, on average, higher near the start of the barrier but stabilizes at longer times. Similar calculations of transient processes for switching from a wide barrier to a structure with narrow, unequal barriers and a quantum well show slower current growth. This is explained by the reflections from the barriers required to form resonant levels in the well. To achieve complete RT, the barrier heights must be sufficiently close. Numerical calculations of the transparency coefficients show not only full resonances but also peaks with incomplete RT, where the maxima $D < 1$. Regarding lifetimes $\tau_n = 2\hbar/E_n''$, they are significantly shorter than the corresponding times determined at short times from transient processes as a result of wave packet evolution. This is because the wave packet contains

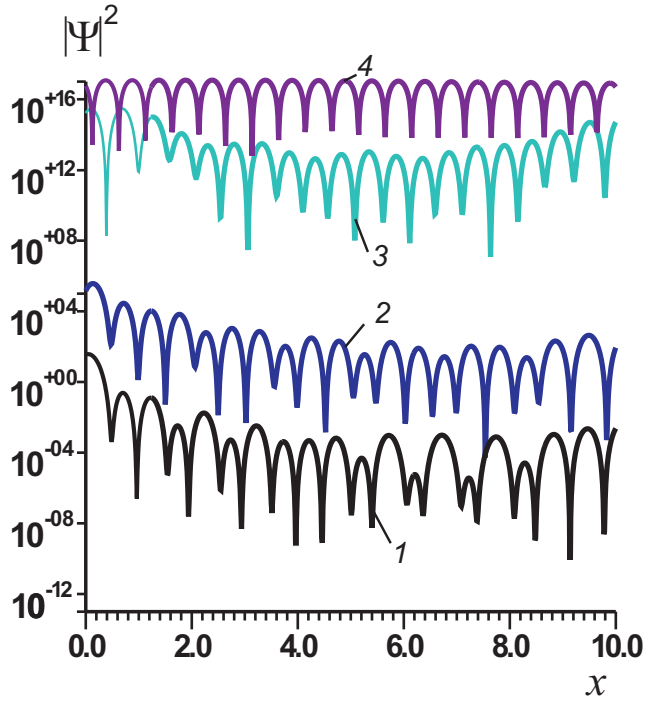

Fig. 7. Particle number density (m^3) as a function of the coordinate x (nm) in a vacuum diode structure at different moments in time (fs): 0.1 (*1*), 0.3 (*2*), 0.5 (*3*), 1.0 (*4*).

a broad energy spectrum. At longer times, the non-exponential nature of the level decay becomes evident (see, for example, [30–34]), with contributions from algebraic terms. Determining level lifetimes this way is feasible only for very narrow wave packets, which is challenging to achieve experimentally for non-relativistic quantum particles, and even more difficult to observe their passage through a barrier. This raises problems with reflecting a spectrally narrow (i.e., spatially very broad) wave packet from the barrier [18], especially when the barrier itself changes over time. However, the macroscopic current density can be measured with high accuracy.

The quantity with the dimension of velocity:

$$v(x,t) = j(x,t)/|\Psi(x,t)|^2$$

can be interpreted as the speed of the probability density movement at point x at time t. This corresponds to the concept introduced by N.A. Umov, but it cannot be interpreted as the speed of an individual particle. For a single-speed particle flow, it coincides with the particle velocity in the flow. The increase in current is accompanied by an increase in the probability density of particle presence inside the barrier. The average instantaneous speed of the wave packet (WP) passing through point x over time τ can be defined as:

$$\bar{v}(x,t,\tau) = \frac{1}{\tau}\int_{t-\tau}^{t} \frac{j(x,t')}{|\Psi(x,t')|^2}dt'. \qquad (18)$$

If the WP is finite in time, its average speed can also be determined.

Short lifetimes of quasi-stationary levels are essential for achieving high current densities in field emission. It is desirable to have as many such levels as possible, and sufficiently deep ones. Increasing the number of levels is achieved by increasing the width of the quantum well, while reducing lifetimes is achieved by using narrow-width barriers. Current growth is also facilitated by leveling the barrier heights, which can be controlled by the gate voltage and the change in the gate work function.

FUNDING

This work was supported by the Ministry of Education and Science of the Russian Federation under the state assignment FSRR-2023-0008 and the Strategic Academic Leadership Program of RUDN University, project 021934-0-000.

APPENDIX

For the well, from the conditions at the cathode-side barrier, we have:

$$A_w^+ = \frac{A_c^+ \exp(-\tilde{k}_c t_c) + A_c^- \exp(\tilde{k}_c t_c)}{2} + \frac{\tilde{k}_c A_c^- \exp(\tilde{k}_c t_c) - \tilde{k}_c A_c^+ \exp(-\tilde{k}_c t_c)}{2ik},$$

$$A_w^- = \frac{A_c^+ \exp(-\tilde{k}_c t_c) + A_c^- \exp(\tilde{k}_c t_c)}{2} - \frac{\tilde{k}_c A_c^- \exp(\tilde{k}_c t_c) - \tilde{k}_c A_c^+ \exp(-\tilde{k}_c t_c)}{2ik}.$$

In the case of wide barriers, neglecting exponentially small terms (reflections from the left edge of the barrier with amplitude A_c^+), we find from the matching conditions at the cathode-side barrier:

$$A_w^+ \approx \frac{A_c^- \exp(\tilde{k}_c t_c)(1 - i\tilde{k}_a/k_0)}{2},$$

$$A_w^- \approx \frac{A_c^- \exp(\tilde{k}_c t_c)(1 + i\tilde{k}_a/k_0)}{2}.$$

On the other hand, the matching conditions at the anode-side barrier give:

$$A_w^+ = \frac{\exp(-ik_0t_w)\left[A_a^+ + A_a^- + i\left(A_a^+ - A_a^-\right)\tilde{k}_a/k_0\right]}{2},$$

$$A_w^- = \frac{\exp(ik_0t_w)\left[A_a^+ + A_a^- - i\left(A_a^+ - A_a^-\right)\tilde{k}_a/k_0\right]}{2}.$$

At the cathode boundary, we have the relations:

$$A_c^+ = \frac{A_c\left(1 + ik_0/\tilde{k}_c\right)}{2},$$

$$A_c^- = \frac{A_c\left(1 - ik_0/\tilde{k}_c\right)}{2}.$$

At the anode boundary, we have accordingly:

$$A_a^+ = \frac{A_a\exp\left(\tilde{k}_at_a\right)\left(1 - ik_a/\tilde{k}_a\right)}{2},$$

$$A_a^- = \frac{A_a\exp\left(-\tilde{k}_at_a\right)\left(1 - ik_a/\tilde{k}_a\right)}{2}.$$

For wide barriers, the amplitudes A_a^- and A_A^+ are small. Assuming them to be zero, we obtain:

$$A_w^+ \approx$$

$$\approx\frac{A_a\exp\left(\tilde{k}_at_a\right)\left(1-ik_0/\tilde{k}_a\right)\left(1+i\tilde{k}_a/k_0\right)\exp\left(-ik_0t_w\right)}{4},$$

$$A_w^+ \approx \frac{A_c\exp\left(\tilde{k}_ct_c\right)\left(1 - ik_0/\tilde{k}_c\right)\left(1 - i\tilde{k}_c/k_0\right)}{4},$$

$$A_w^- =$$

$$=\frac{A_a\exp\left(\tilde{k}_at_a\right)\left(1-ik_0/\tilde{k}_a\right)\left(1-i\tilde{k}_a/k_0\right)\exp\left(ik_0t_w\right)}{4},$$

$$A_w^- \approx \frac{A_c\exp\left(\tilde{k}_ct_c\right)A_c\left(1 - ik_0/\tilde{k}_c\right)\left(1 + i\tilde{k}_c/k_0\right)}{4}.$$

Equating the coefficients $A_w^{\pm}$, we get two equations:

$$A_a\exp\left(\tilde{k}_at_a\right)\left(1-ik_0/\tilde{k}_a\right)\left(1+i\tilde{k}_a/k_0\right)\exp\left(-ik_0t_w\right) =$$

$$= A_c\exp\left(\tilde{k}_ct_c\right)\left(1 - ik_0/\tilde{k}_c\right)\left(1 - i\tilde{k}_c/k_0\right),$$

$$A_a\exp\left(\tilde{k}_at_a\right)\left(1 - ik_0/\tilde{k}_a\right)\left(1 - i\tilde{k}_a/k_0\right)\exp\left(ik_0t_w\right) =$$

$$= A_c\exp\left(\tilde{k}_ct_c\right)\left(1 - ik_0/\tilde{k}_c\right)\left(1 + i\tilde{k}_c/k_0\right).$$

Dividing the first by the second, we obtain the approximate characteristic equation:

$$\frac{\left(1 - i\tilde{k}_A/k_0\right)\left(1 - i\tilde{k}_a/k_0\right)}{\left(1 + i\tilde{k}_A/k_0\right)\left(1 + i\tilde{k}_a/k_0\right)} = \exp\left(-2ik_0t_w\right). \qquad \text{(A1)}$$

To obtain the exact equation, all amplitudes must be retained. In this case, equating the coefficients gives:

$$\begin{pmatrix} A_A^+ \\ A_A^- \end{pmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}\begin{pmatrix} A_a^+ \\ A_a^- \end{pmatrix},$$

$$\begin{pmatrix} A_a^+ \\ A_a^- \end{pmatrix} = \begin{bmatrix} M_{11}^{-1} & M_{12}^{-1} \\ M_{21}^{-1} & M_{22}^{-1} \end{bmatrix}\begin{pmatrix} A_c^+ \\ A_c^- \end{pmatrix}.$$

The matrix elements $\widehat{M}$ are given by:

$$M_{11} = \exp\left(\tilde{k}_At_c\right)\tilde{M}_{11} = \exp\left(\tilde{k}_At_c\right)\times$$

$$\times\frac{\cos\left(k_0t_w\right)\left(1+\tilde{k}_a/\tilde{k}_c\right)+\left(\tilde{k}_a/k_0-k_0/\tilde{k}_A\right)\sin\left(kt_w\right)}{2},$$

$$M_{12} = \exp\left(\tilde{k}_At_c\right)\tilde{M}_{12} = \exp\left(\tilde{k}_At_c\right)\times$$

$$\times\frac{\cos\left(k_0t_w\right)\left(1-\tilde{k}_a/\tilde{k}_A\right)-\left(\tilde{k}_a/k_0+k_0/\tilde{k}_A\right)\sin\left(kt_w\right)}{2},$$

$$M_{21} = \exp\left(-\tilde{k}_At_c\right)\tilde{M}_{21} = \exp\left(-\tilde{k}_At_c\right)\times$$

$$\times\frac{\cos\left(k_0t_w\right)\left[1-\tilde{k}_a/\tilde{k}_A\right]+\left(\tilde{k}_a/k_0+k_0/\tilde{k}_c\right)\sin\left(k_0t_w\right)}{2},$$

$$M_{22} = \exp\left(-\tilde{k}_At_c\right)\tilde{M}_{22} = \exp\left(-\tilde{k}_At_c\right)\times$$

$$\times\frac{\cos\left(k_0t_w\right)\left(1+\tilde{k}_a/\tilde{k}_c\right)-\left(\tilde{k}_a/k_0-k_0/\tilde{k}_c\right)\sin\left(k_0t_w\right)}{2}.$$

Now

$$A_c^+ = A_c\left(1 + ik_0/\tilde{k}_c\right) =$$

$$= 2\left(M_{11}A_a^+ + M_{12}A_a^-\right),$$

$$A_c^- = A_c\left(1 - ik_0/\tilde{k}_c\right) =$$

$$= 2\left(M_{21}A_a^+ + M_{22}A_a^-\right).$$

Substituting $A_a^{\pm}$ in these expressions, we obtain:

$$A_c\left(1+ik_0/\tilde{k}_c\right)=M_{11}A_a\exp\left(\tilde{k}_at_a\right)\left(1-ik_0/\tilde{k}_a\right)+$$

$$+M_{12}A_a\exp\left(-\tilde{k}_at_a\right)\left(1 - ik_0/\tilde{k}_a\right),$$

$$A_c\left(1-ik_0/\tilde{k}_c\right)=M_{21}A_a\exp\left(\tilde{k}_a t_a\right)\left(1-ik_0/\tilde{k}_a\right)+$$
$$+M_{22}A_a\exp\left(-\tilde{k}_a t_a\right)\left(1-ik_0/\tilde{k}_a\right).$$

Dividing the first equation by the second, we obtain the characteristic equation:

$$\frac{1+ik_0/\tilde{k}_c}{1-ik_0/\tilde{k}_c}=f(E)= \tag{A2}$$
$$=\frac{\tilde{M}_{11}\exp\left(\tilde{k}_c t_c+\tilde{k}_a t_a\right)+\tilde{M}_{12}\exp\left(\tilde{k}_c t_c-\tilde{k}_a t_a\right)}{\tilde{M}_{21}\exp\left(\tilde{k}_a t_a-\tilde{k}_c t_c\right)+\tilde{M}_{22}\exp\left(-\tilde{k}_a t_a-\tilde{k}_c t_c\right)}.$$

For wide barriers, small terms can be neglected, resulting in the simplified form:

$$f(E)\approx\exp\left(2\tilde{k}_c t_c\right)\times$$
$$\times\frac{1+\tilde{k}_a/\tilde{k}_c+\left(\tilde{k}_a/k_0-k_0/\tilde{k}_c\right)\left(k_0 t_w\right)}{1-\tilde{k}_a/\tilde{k}_c+\left(\tilde{k}_a/k_0+k_0/\tilde{k}_c\right)\left(k_0 t_w\right)}.$$

The matrix elements appearing in equations (14) and (15) are expressed as:

$$A_{nm}^{cc}=\frac{sinc\left(\left(\chi_n-\chi_m\right)d\right)+sinc\left(\left(\chi_n+\chi_m\right)d\right)}{1+sinc\left(2\chi_n d\right)}=$$
$$=\frac{\delta_{nm}}{1+\delta_{n0}},$$
$$A_{nm}^{cs}=\frac{conc\left(\left(\chi_n-\chi_m\right)d\right)-conc\left(\left(\chi_n+\chi_m\right)d\right)}{1+sinc\left(2\chi_n d\right)}=$$
$$=\frac{(-1)^{n+m}-(-1)^{n-m}}{1+\delta_{n0}}, \tag{A3}$$
$$A_{nm}^{sc}=\frac{conc\left(\left(\chi_n-\chi_m\right)d\right)+conc\left(\left(\chi_n+\chi_m\right)d\right)}{1+sinc\left(2\chi_n d\right)}=$$
$$=\frac{2-(-1)^{n-m}-(-1)^{n+m}}{1+\delta_{n0}},$$
$$A_{nm}^{ss}=\frac{sinc\left(\left(\chi_n-\chi_m\right)d\right)-sinc\left(\left(\chi_n+\chi_m\right)d\right)}{1-sinc\left(2\chi_n d\right)}=\delta_{nm}.$$

These expressions involve the following integrals:

$$V_{nm}^{\alpha cc}(t)=\frac{2}{(1+\delta_{n0})d}\int_0^d V(t,x)\cos\left(\chi_n x\right)\cos\left(\chi_m x\right)dx,$$
$$V_{nm}^{\beta ss}(t)=\frac{2}{(1+\delta_{n0})d}\int_0^d V(t,x)\cos\left(\chi_n x\right)\sin\left(\chi_m x\right)dx,$$
$$V_{nm}^{\beta ss}(t)=2\int_0^d V(t,x)\sin\left(\chi_n x\right)\sin\left(\chi_m x\right)dx,$$
$$V_{nm}^{\alpha sc}(t)=2\int_0^d V(t,x)\sin\left(\chi_n x\right)\cos\left(\chi_m x\right)dx.$$

In equation (A3), the functions $sinc(x)=\sin(x)/x$ and $conc(x)=(1-\cos(x))/x$ are included. These functions at zero should be defined as $sinc(0)=1$, $conc(0)=0$ ensuring proper boundary conditions. Moreover, the condition $sinc(2n\pi)=\delta_{n0}$ applies at the barrier edge.

REFERENCES

1. M.V. Davidovich, I.S. Nefedov, O.E. Glukhova, M.M. Slepchenkov, J. Appl. Phys. **130**(20), 204301 (2021).
2. M.V. Davidovich, Technical Physics **67**(9), 1196 (2022).
3. M.V. Davidovich, I.S. Nefedov, O.E. Glukhova, M.M. Slepchenkov, J.M. Rubi, Scientific Reports **13**, 19365 (2023).
4. R. Tsu, L. Esaki, Appl. Phys. Lett. **22**(11), 562 (1973).
5. L.L. Chang, L. Esaki, R. Tsu, Appl. Phys. Lett. **24**, 593 (1974).
6. E.X. Ping, H.X. Jiang, Phys. Rev. B **40**(17), 11792 (1989).
7. O. Pinaud, J. Appl. Phys. **92**(4), 1987 (2002).
8. V.F. Elesin, Journal of Experimental and Theoretical Physics **101**(5), 795 (2005).
9. V.F. Elesin, Journal of Experimental and Theoretical Physics **118**(6), 951 (2014).
10. V.F. Elesin, Yu.V. Kopaev, Journal of Experimental and Theoretical Physics **96**(6), 1149 (2003).
11. K.S. Grishakov, V.F. Elesin, Semiconductors **50**(8), 1092 (2016).
12. M.V. Davidovich, JETP Letters **110**(7), 472 (2019).
13. E.A. Nelin, Phys. Usp. **50** (3), 293 (2007).
14. J.G. Simmons, J. Appl. Phys. **34** (6), 1793 (1963).
15. E.R. Lubenets, Theoretical and Mathematical Physics **32**(2), 741 (1977).
16. G.F. Drukarev, Sov. Phys. JETP **51**(1), 59 (1951).
17. A.I. Baz', Ya.B. Zel'dovich, A.M. Perelomov, Scattering, reactions and decays in nonrelativistic quantum mechanics (Nauka, Moscow, 1971). [In Russian].
18. M.V. Davidovich, JETP **130**(1), 35 (2020).
19. A.S. Davydov, Quantum Mechanics (Pergamon, 1965).
20. L.D. Landau, E.M. Lifshitz, Quantum Mechanics. Nonrelativistic Theory (Pergamon, 1965).

21. V.N. Gribov, Quantum electrodynamics (SIC Regular and chaotic dynamics, Izhevsk, 2001). [In Russian].

22. A.N. Tikhonov, A.A. Samarsky, Equations of Mathematical Physics (Nauka, Moscow 1977). [In Russian].

23. L. Fonda, G.C. Ghirardi, A. Rimini, Rep. Prog. Phys. **41**(4), 587 (1978).

24. E.C.G. Sudarshan, C.B. Chiu, G. Bhamathi, Adv. Chem. Phys. **99**(3), 121 (1997).

25. L.M. Krauss, J. Dent, Phys. Rev. Lett. **100**(17), 171301 (2008).

26. L.A. Khalhin, JETP **33**(6), 1371 (1958).

27. C.B. Chiu, E.C.G. Sudarshan, B. Misra, Phys. Rev. D **16**(2), 520 (1977).

28. G. García-Calderón, R. Romo, Phys. Rev. A **100**(3), 032121 (2019).

29. P. Facchi, S. Pascazio, J. Phys. A **41**(49), 493001, (2008).

30. C. Rothe, S.I. Hintschich, A.P. Monkman, Phys. Rev. Lett. **96**(16), 163601 (2006).

31. C. Anastopoulos, Int. J. Theor. Phys. **58**(3), 890 (2019).

32. S.D. Druger, M.A. Samuel, Phys. Rev. A **30**(1), 640 (1984).

33. C.A. Nicolaides, Phys. Rev. A **66**(2), 022118 (2002).

34. N.G. Kelkar, M. Nowakowski, K.P. Khemchandani, Phys. Rev. C **70**(2), 024601 (2004).

35. R.G. Newton, Ann. Phys. **14**(1), 333 (1961).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 24–41

ATOMS, MOLECULES, OPTICS

XUV-INDUCED CHANNELS FOR STIMULATED GENERATION OF RADIATION BY AN ATOM IN AN INTENSE IR LASER FIELD

© 2025 Ya. V. Breev[a, b], A. A. Minina[a, b], A. V. Flegel[a, b,*], M. V. Frolov[a, b,**]

[a]Department of Physics, Voronezh State University, Voronezh, Russia
[b]Department of Radiophysics, Lobachevsky State University, Nizhny Novgorod, Russia
*e-mail: flegel@cs.vsu.ru
**e-mail: frolov@phys.vsu.ru

Received September 13, 2024

Revised September 13, 2024

Accepted September 26, 2024

Abstract. The perturbation theory in interaction of isolated attosecond XUV pulse with an atomic system subjected to an intense IR field has been developed. Analytical expressions have been obtained for XUV-induced corrections to the wave function and amplitude of radiation generation in arbitrary order of perturbation theory. The contribution of partial amplitudes of radiation generation is analyzed for different channels with absorption of XUV photons at the ionization and recombination steps of the rescattering scenario. The region of **IR- and XUV-pulse** parameters are identified, in which interference of different XUV-induced channels is possible up to the third order of perturbation theory in XUV interaction.

Keywords: *isolated attosecond XUV-pulse, intense IR field, adiabatic approach, perturbation theory, high harmonic generation, rescattring effects*

DOI: 10.31857/S00444510250102e4

1. INTRODUCTION

Significant progress has been made in the theoretical description of nonlinear effects arising from the interaction of intense infrared (IR) laser fields with atomic systems. The quantum mechanical description of processes induced by an intense IR field involves two approaches: numerical methods, such as solving the time-dependent Schrödinger equation (TDSE) and its simplified variations for multi-electron systems (density functional theory, time-dependent Hartree-Fock method) [1–9], and analytical approaches. Numerical calculations typically serve as "benchmarks" for verifying the accuracy of analytical approaches and demonstrate their efficiency in determining the nonlinear response of an atomic system to an intense external alternating electric field. However, the results of numerical integration can only be obtained for fixed laser parameters and lack significant predictive power. Specifically, in most cases, it is necessary to perform numerous time-consuming computations to achieve the desired physical interpretation of the observed effect. In contrast, analytical theories are better suited for uncovering general fundamental patterns in the nonlinear interaction of an atomic system with an intense laser field.

Analytical approaches to describing nonlinear effects in the interaction of IR fields with atomic or molecular systems are typically based on the single-electron approximation. Within this approximation (subject to certain obvious limitations), it becomes possible to derive expressions for the amplitudes and cross sections of fundamental laser-induced and laser-assisted atomic processes with accuracy not inferior to numerical results of TDSE solution [10–16].

A key advantage of analytical approaches over numerical methods is the ability to establish a universal parameterization dependence of the probabilities of the strong-field processes on the fundamental characteristics of the target (i.e. the electron-core interaction potential $U(r)$) and the laser-pulse parameters (see, e.g., [17]).

These parameterizations can be further generalized to multi-electron systems, enabling the study of

the effects of internal electronic dynamics in laser-induced photoprocesses [18].

Among the analytical approaches, the most popular is the S-matrix formalism, where the exact wave function of the active electron in the self-consistent potential $U(r)$ is expanded into a formal series in $U(r)$ [19, 20] (see also [21, 22]). This expansion leads to a Born series for the transition amplitude, where the terms (partial transition amplitudes) can be expressed as a convolution of the free-electron Green's function in the laser field with the atomic potential. For example, for the above-threshold ionization (ATI) process, the account of $U(r)$ in the lowest order leads to the Keldysh result [23].

Due to the large value of the classical action of the electron in a strong low-frequency field, the partial amplitudes can be analyzed using the saddle-point method [24], which gives rise to the quantum orbit approach [25, 26]. This approach provides an intuitive physical interpretation of strong-field phenomena in terms of classical trajectories, thereby justifying the rescattering model for fundamental atomic photoprocesses in an intense laser field [19, 20, 27, 28].

Although the Born expansion of transition amplitudes has proven fruitful and significantly contributes to the description of strong-field phenomena, it cannot fully account for the atomic potential, whose influence can be crucial [9, 18, 29–32]. One approach that allows for a more accurate treatment of the atomic system dynamics in an intense low-frequency field is the adiabatic approximation. The general idea of this approximation is based on the smallness of the carrier frequency ω of the laser pulse compared to the ionization threshold I_p of the atomic target $\hbar\omega \ll I_p$. The lowest-order of the adiabatic approximation (zero-order approximation) is defined by the quasistationary state of the system in a static (DC) field with an intensity equal to the instantaneous value of the low-frequency laser field [33–37]. In [11–17], a correction to the zero-order adiabatic approximation was derived for the wave function, accounting for the rescattering of the electron by the atomic potential. The study [38] refined the adiabatic approach for determining the atomic state in the lowest adiabatic approximation, by utilizing the analytical part of the wave function of the quasistationary atomic state in the instantaneous laser field. Within the adiabatic approach, both low-energy and high-energy (rescattering plateaus) parts of the photoelectron spectra and high-harmonic generation (HHG) spectra have been calculated.

The presence of a closed analytical expression for the wave function of the atomic state in an intense IR field allows for the development of an adiabatic perturbation theory in additional interaction with a high-frequency (e.g., extreme ultraviolet – XUV) attosecond pulse [17, 39]. The influence of an ultrashort XUV pulse on the radiation generation process results in the appearance of a significant number of new generation channels and substantial modification of the IR field HHG spectra. For example, the enhancement of harmonic yield due to the resonant population of excited target states by the XUV pulse was studied in [40–43]. XUV-induced enhancement of high harmonic yield was investigated both for attosecond pulse train [44–47] and for an isolated attosecond XUV pulse [48, 49]. These studies demonstrated that a XUV pulse (or its sequence) can affect the ionization stage in the three-step Corkum model [50], i.e., change the ionization times and thereby affect the harmonic yield. In [51, 52], it was shown that adding a weak XUV field leads to the appearance of an additional plateau in the HHG spectra. The physics of the additional XUV-induced plateau was explained in [53], where it was shown that the additional plateau results from XUV-photon absorption at the recombination stage. It should also be noted that, at sufficiently high carrier frequencies of the XUV pulse, electrons from the inner atomic shell can also participate in the HHG process, leading to an increase in the cutoff energy of the plateau [54–56]. Moreover, such XUV pulses, combined with an intense IR field, enable the study of Auger processes [57, 58] and electronic transitions from inner shells to the valence shell [59]. The re-emission channel (or elastic scattering) of a XUV photon by the atomic system, leading to significant enhancement of the generated radiation yield, was studied in [60]. Second-order processes of XUV interaction in an IR-dressed atomic medium were also investigated: generation of a XUV pulse at the doubled carrier frequency [39] and the XUV pulse rectification effect [61].

In this paper, we generalize the perturbative approach proposed in [17, 39] to construct perturbation theory corrections for the interaction

with a short XUV pulse of arbitrary order, based on adiabatic wave functions of the atomic state in an intense IR field. Within the developed perturbation theory, XUV-induced radiation generation channels are investigated by analyzing classical electron trajectories in the field of synchronized intense IR and attosecond XUV pulses, and the possibility of interference between different channels due to their spectral overlap is explored. The article uses atomic units unless otherwise specified.

2. ADIABATIC APPROACH TO THE DESCRIPTION OF AN ATOM IN A LOW-FREQUENCY LASER FIELD

2.1. Adiabatic expression for the wave function

Let us consider the interaction of an atomic system with an intense infrared (IR) laser pulse characterized by the peak electric field strength F_{IR} and the carrier frequency ω_{IR}. We will assume that the laser pulse parameters satisfy the adiabatic conditions [23]:

$$\omega_{IR} \ll | E_0 |, \quad \gamma_K \ll 1, \tag{1}$$

where $\gamma_K = \kappa\omega_{IR} / F_{IR}$ is the Keldysh parameter, $\kappa = \sqrt{2 | E_0 |}$, E_0 is related to the binding energy of the unperturbed atomic level. The conditions (1) can also be rewritten in terms of the average oscillation energy of a free electron in the laser field $u_p = F_{IR}^2 / (4\omega_{IR}^2)$:

$$\omega_{IR} \ll | E_0 |, \quad \omega_{IR} \ll u_p. \tag{2}$$

To describe the nonlinear interaction of an atomic system with a laser field that satisfies the conditions (2), it is most convenient to use the adiabatic approach [12, 13, 37]. Within this approach, the wave function of an atomic electron interacting with a low-frequency laser field can be represented as a sum of "slow" ($\Psi_{IR}^{(0)}(\mathbf{r},t)$) and "fast" ($\Psi_{IR}^{(r)}(\mathbf{r},t)$) time-dependent parts [11, 12, 17]:

$$\Psi_{IR}(\mathbf{r},t) = \Psi_{IR}^{(0)}(\mathbf{r},t) + \Psi_{IR}^{(r)}(\mathbf{r},t). \tag{3}$$

The slow part $\Psi_{IR}^{(0)}(\mathbf{r},t)$ represents the adiabatic approximation in the lowest order ("zero-order" approximation) and is defined by the quasistationary state in a DC electric field with a strength equal to the instantaneous value of the IR field at time t [37, 12]. In many practical calculations, the function $\Psi_{IR}^{(0)}(\mathbf{r},t)$ can be accurately approximated by the initial-state wave function in the absence of the IR field:

$$\Psi_{IR}^{(0)}(\mathbf{r},t) \approx e^{-iE_0 t}\varphi_0(\mathbf{r}). \tag{4}$$

The term $\Psi_{IR}^{(r)}(\mathbf{r},t)$ in Eq. (3) describes the rescattering effects of the valence electron on the atomic core and represents a superposition of scattering states $\psi_{\mathbf{K}_s}^{(+)}$ of the electron in the atomic potential with laser-induced momenta $\mathbf{K}_s$ [17]:

$$\Psi_{IR}^{(r)}(\mathbf{r},t) = e^{-iE_0 t}\Phi_{IR}^{(r)}(\mathbf{r},t), \tag{5a}$$

$$\Phi_{IR}^{(r)}(\mathbf{r},t) = \sum_s a_s(t)\psi_{\mathbf{K}_s}^{(+)}(\mathbf{r}). \tag{5b}$$

Each term in the sum (5b) is associated with one of the possible closed classical trajectories, which start at the tunneling time t_s' and end at the return time t of the electron back to the atomic core. The laser-induced momenta are defined by the expression:

$$\mathbf{K}_s = \mathbf{K}(t,t_s'), \tag{6}$$

$$\mathbf{K}(t,t') = \mathbf{A}_{IR}(t) - \frac{1}{t-t'}\int_{t'}^{t}\mathbf{A}_{IR}(\tau)d\tau,$$

where $\mathbf{A}_{IR}(t)$ is the vector potential associated with the electric field strength $\mathbf{F}_{IR}(t)$ of the laser pulse by the relation:

$$\mathbf{F}_{IR}(t) = -\partial\mathbf{A}_{IR}(t) / \partial t.$$

The tunneling times $t_s'(t)$ as functions of the return times t satisfy the transcendental equation (see details in [14]):

$$\mathbf{K}_s' \cdot \dot{\mathbf{K}}_s' = 0, \tag{7}$$

where

$$\mathbf{K}_s' \equiv \mathbf{K}'(t,t_s'),$$
$$\dot{\mathbf{K}}_s' = \frac{\partial\mathbf{K}_s'}{\partial t_s'}, \tag{8}$$
$$\mathbf{K}'(t,t') = \mathbf{A}_{IR}(t') - \frac{1}{t-t'}\int_{t'}^{t}\mathbf{A}_{IR}(\tau)d\tau.$$

Equation (7) has a simple physical meaning: the atomic electron tunnels at the moments in time that provide the minimum kinetic energy of the released electron in the laser field. The time-dependent coefficients $a_s(t)$ in the superposition (5b) represent

the product of the ionization (tunneling) factor $a^{(tun)}(t'_s)$ and the propagation factor $a^{(pr)}(t,t'_s)$:

$$a_s(t) = a^{(tun)}(t'_s)a^{(pr)}(t,t'_s). \quad (9)$$

The ionization factor is characterized by the tunneling exponent in the instantaneous "static" electric field with the strength:

$$\mathcal{F} = [\mathbf{F}^2_{IR}(t'_s) - \mathbf{K}'_s \cdot \dot{\mathbf{F}}_{IR}(t'_s)]^{1/2},$$

see [62]. For example, in the case of a linearly polarized laser field, the following relation holds:

$$a^{(tun)}(t'_s) \propto e^{-F_{at}/(3|F_{IR}(t'_s)|)}, \quad (10)$$

where $F_{at} = \kappa^3$ defines the magnitude of the characteristic intra-atomic field. To satisfy the quasiclassical condition, an additional inequality must hold: $F \ll F_{at}$, which ensures the smallness of the ionization factor and the insignificance of the initial-state decay effects.

The propagation factor $a^{(pr)}(t,t'_s)$ is determined by the classical action $S(t,t'_s)$ of the free electron in the laser field over the time interval from t'_s to t:

$$a^{(pr)}(t,t'_s) = \frac{e^{iE_0(t-t'_s)-iS(t,t'_s)}}{(t-t'_s)^{3/2}}, \quad (11a)$$

$$S(t,t'_s) = \frac{1}{2}\int_{t'_s}^{t}\Big[\mathbf{A}_{IR}(\tau) - \frac{1}{t-t'_s}\int_{t'_s}^{t}\mathbf{A}_{IR}(\tau')d\tau'\Big]^2 d\tau. \quad (11b)$$

It is important to note that the rescattered part $\Phi^{(r)}_{IR}(\mathbf{r},t)$ of the atomic electron wave function in the IR field, relative to the unperturbed function $\varphi_0(\mathbf{r})$, has a smallness $\sim \beta_{IR}$

$$\beta_{IR} = \gamma_K^{3/2}\frac{F_{IR}}{F_{at}}e^{-F_{at}/(3F_{IR})} \ll 1.$$

Essentially, the result (3) represents an expansion of the atomic electron state in terms of β_{IR} up to the first order. In the following, we will maintain this accuracy, as the inclusion of higher-order terms in the expansion of β_{IR} (i.e., a more precise account of rescattering effects) does not lead to any significant manifestations in the amplitudes and cross sections of processes in a strong IR laser field.

2.2 Amplitude of radiation generation

The amplitude of photon generation by an atom in an intense laser field is determined by the dipole matrix element [63, 64]:

$$\mathbf{D}(\Omega) = \int\langle\tilde{\Psi}_{IR}(\mathbf{r},t)|\mathbf{r}|\Psi_{IR}(\mathbf{r},t)\rangle e^{i\Omega t}dt, \quad (12)$$

where Ω is the frequency of the generated photon, $\tilde{\Psi}_{IR}(\mathbf{r},t)$ is the dual wave function to the state $\Psi_{IR}(\mathbf{r},t)$, defined from the state $\Psi_{IR}(\mathbf{r},t)$ by complex conjugation, time reversal $t \to -t$, and the replacement of all t-odd parameters λ for $-\lambda$ [65, 66]. In a low-frequency laser field, the dipole matrix element (12) for $\Omega > |E_0|$ can be approximately expressed through $\Psi^{(0)}(\mathbf{r},t)$ and $\Psi^{(r)}_{IR}(\mathbf{r},t)$ [17, 63]:

$$\mathbf{D}(\Omega) = \int\langle\varphi_0 e^{-iE_0 t}\,|\,\mathbf{r}\,|\,\Psi^{(r)}_{IR}(\mathbf{r},t)\rangle e^{i\Omega t}dt. \quad (13)$$

The harmonic yield, summed over polarizations and integrated over directions, is determined by the square of the modulus of $\mathbf{D}(\Omega)$:

$$Y = \frac{\Omega^4|\mathbf{D}(\Omega)|^2}{4\pi^2 c^3},$$

where c is the speed of light.

In the adiabatic approximation, the time integral in (13) is evaluated using the saddle-point method, and $\mathbf{D}(\Omega)$ can be represented as a sum of partial amplitudes $\mathbf{D}_j(\Omega)$ [14, 17]:

$$\mathbf{D}(\Omega) = \sum_j \mathbf{D}_j(\Omega), \quad (14a)$$

$$\mathbf{D}_j(\Omega) = a_j^{(tun)}\bar{a}_j\mathbf{d}(\mathbf{K}_j)e^{i\Omega t_j}, \quad (14b)$$

where $a_j^{(tun)}$, $\bar{a}_j$ are the tunneling and propagation factors, respectively, and $\mathbf{d}(\mathbf{K}_j)$ is the dipole matrix element for the transition from the continuum state with momentum $\mathbf{K}_j$ to the bound state $\varphi_0(\mathbf{r})$:

$$\mathbf{d}(\mathbf{K}_j) = \langle\varphi_0(\mathbf{r})\,|\,\mathbf{r}\,|\,\psi^{(+)}_{\mathbf{K}_j}(\mathbf{r})\rangle.$$

The factors $a_j^{(tun)}$, $\bar{a}_j$ are defined by the relations:

$$a_j^{(tun)} \equiv a^{(tun)}(t'_j),$$

$$\bar{a}_j = \sqrt{\frac{2\pi i}{\mathbf{K}_j\cdot\mathbf{F}_{IR}(t_j) + \dfrac{\mathbf{K}_j^2}{t_j - t'_j}}}\,a^{(pr)}(t_j,t'_j).$$

The summation in (14a) is performed over all closed classical electron trajectories, defined by the start time t'_j and end time t_j of the electron's motion. The times t'_j and t_j are the roots of the system of transcendental equations [14, 17]:

$$\mathbf{K}'_j \cdot \dot{\mathbf{K}}'_j = 0, \quad \mathbf{K}'^2_j = 2(\Omega + E_0), \tag{15}$$

where the induced momenta $\mathbf{K}'_j \equiv \mathbf{K}'(t_j,t'_j)$, $\mathbf{K}_j \equiv \mathbf{K}(t_j,t'_j)$ are defined in Eqs. (8) and (6), respectively.

3. TIME-DEPENDENT PERTURBATION THEORY FOR AN ATOMIC SYSTEM IN AN INTENSE IR FIELD

Let us consider an atomic system interacting with an intense IR field and a perturbative XUV pulse. The account of the XUV interaction with the IR-dressed atomic system can be treated within the perturbation theory based on the adiabatic wave functions of the atomic electron in the IR field [17]. We will consider the interaction with the XUV pulse in the dipole approximation, so that the potential $V_{XUV}(\mathbf{r},t)$ of the interaction between the atomic electron and the XUV pulse has the form:

$$V_{XUV}(\mathbf{r},t) = V_+(\mathbf{r},t)e^{-i\omega_{XUV}t} + V_-(\mathbf{r},t)e^{i\omega_{XUV}t}, \tag{16}$$

$$V_+(\mathbf{r},t) = \frac{F_{XUV}}{2}(\mathbf{e}_{XUV}\cdot\mathbf{r})f_{XUV}(t),$$

$$V_-(\mathbf{r},t) = V^*_+(\mathbf{r},t),$$

where F_{XUV} is the peak field strength, ω_{XUV} is the carrier frequency, $\mathbf{e}_{XUV}$ is the polarization vector, and $f_{XUV}(t)$ is the XUV pulse envelope.

Note that for $\omega_{XUV} > |E_0|$, the small perturbation parameter for the XUV interaction is defined as [67]

$$\beta_{XUV} = \frac{\kappa F_{XUV}}{\omega^2_{XUV}} = 4\frac{|E_0|^2}{\omega^2_{XUV}}\frac{F_{XUV}}{F_{at}} \ll 1. \tag{17}$$

Therefore, even in the case of XUV radiation strength comparable to F_{at}, the interaction V_{XUV} can be treated perturbatively [68].

The state $\Psi(\mathbf{r},t)$ of the atomic electron in the field of synchronized IR and XUV pulses can be written as:

$$\Psi(\mathbf{r},t) = \Psi_{IR}(\mathbf{r},t) + \\ + \iint \mathcal{G}(\mathbf{r},t;\mathbf{r}',t')V_{XUV}(\mathbf{r}',t')\Psi_{IR}(\mathbf{r}',t')d\mathbf{r}'dt', \tag{18}$$

where $\mathcal{G}(\mathbf{r},t;\mathbf{r}',t')$ is the time-dependent (retarded) Green's function of the atomic electron in the two-component field. For the function $\mathcal{G}(\mathbf{r},t;\mathbf{r}',t')$, the Dyson equation holds:

$$\mathcal{G}(\mathbf{r},t;\mathbf{r}',t') = \mathcal{G}_{IR}(\mathbf{r},t;\mathbf{r}',t') + \iint \mathcal{G}_{IR}(\mathbf{r},t;\mathbf{r}'',t'')\times \\ \times V_{XUV}(\mathbf{r}'',t'')\mathcal{G}(\mathbf{r}'',t'';\mathbf{r}',t')d\mathbf{r}''dt'', \tag{19}$$

where $\mathcal{G}_{IR}(\mathbf{r},t;\mathbf{r}',t')$ is the time-dependent (retarded) Green's function of the atomic electron in the IR field. Using the relations (18) and (19), we can represent the wave function $\Psi(\mathbf{r},t)$ as a perturbation series in V_{XUV} :

$$\Psi(\mathbf{r},t) = \Psi_0(\mathbf{r},t) + \sum_{n=1}^{\infty}\Psi_n(\mathbf{r},t), \tag{20}$$

where $\Psi_0(\mathbf{r},t) \equiv \Psi_{IR}(\mathbf{r},t)$ is the atomic state in the absence of the XUV pulse (see Eq. (3)), and $\Psi_n(\mathbf{r},t) \sim \beta^n_{XUV}$ are the n-order corrections, satisfying the following recursive relation:

$$\Psi_{n+1}(\mathbf{r},t) = \iint \mathcal{G}_{IR}(\mathbf{r},t;\mathbf{r}',t')\times \\ \times V_{XUV}(\mathbf{r}',t')\Psi_n(\mathbf{r}',t')d\mathbf{r}'dt'. \tag{21}$$

The accuracy of the adiabatic approximation allows for the approximate evaluation of the time integrals in (21). The main contribution to the value of the corresponding integrals are given primarily by the weakly overlapping neighborhoods of the points $t' = t$ and $t' = \tilde{t}'_s(t)$, where $\tilde{t}'_s(t)$ are the saddle points of the phase of the rapidly oscillating factor of the integrand in (21). This phase is primarily determined by the classical action of the electron in the IR field and the carrier frequency of the XUV pulse. The asymptotic expression for the Green's function $\mathcal{G}_{IR}(\mathbf{r},t;\mathbf{r}',t')$ in the vicinity of these singular points was obtained in [17]:

$$\mathcal{G}_{IR}(\mathbf{r},t;\mathbf{r}',t') \approx \\ \approx \begin{cases} G_{at}(\mathbf{r},t;\mathbf{r}',t'), & t \approx t', \\ G_{vol}(0,t;0,t')\psi^{(+)}_{\mathbf{K}}(\mathbf{r})[\psi^{(+)}_{\mathbf{K}'}(\mathbf{r}')]^*, & t \neq t', \end{cases} \tag{22}$$

Where $G_{at}(\mathbf{r},t;\mathbf{r}',t')$ is the time-dependent atomic Green's function of the electron, $G_{vol}(0,t;0,t')$ is the Volkov Green's function of the electron in the IR field for $\mathbf{r} = \mathbf{r}' = 0$, and the momenta $\mathbf{K} \equiv \mathbf{K}(t,t')$ and $\mathbf{K}' \equiv \mathbf{K}'(t,t')$ are defined by Eq. (6) and (8), respectively.

The explicit expression for the nth-order correction $\Psi_n(\mathbf{r},t)$ can be obtained by sequentially calculating the lower-order corrections, using the relations (22) and (21) and approximately evaluating the time integrals. The nth-order correction can be decomposed into a slow part $\Psi_n^{(s)}$ and a fast part $\Psi_n^{(r)}$:

$$\Psi_n(\mathbf{r},t') = \Psi_n^{(s)}(\mathbf{r},t') + \Psi_n^{(r)}(\mathbf{r},t'). \tag{23}$$

To determine the slow part $\Psi_n^{(s)}$, we represent it as a superposition:

$$\Psi_n^{(s)}(\mathbf{r},t) \approx \sum_{\nu} e^{-iE_\nu t}\varphi_\nu^{(n)}(\mathbf{r},t), \tag{24}$$

where $E_\nu = E_0 + \nu\omega_{XUV}$ are time-dependent coefficients, and the slow (in time) functions $\varphi_\nu^{(n)}(\mathbf{r},t)$ require further definition. The slow time dependence of the functions $\varphi_\nu^{(n)}(\mathbf{r},t)$, as well as the envelope $f_{XUV}(t)$, will be understood under the following conditions:

$$\left|\frac{\partial\varphi_\nu^{(n)}}{\partial t}\right| \ll \omega_{XUV}\left|\varphi_\nu^{(n)}\right|, \tag{25a}$$

$$\left|\frac{\partial f_{XUV}(t)}{\partial t}\right| \ll \omega_{XUV}\left|f_{XUV}(t)\right|. \tag{25b}$$

It should be emphasized that in order to isolate the slowly varying part of the wave function, one should neglect the contribution from the saddle point neighborhoods $\tilde{t}_s'(t)$ in the time integral in (21) and consider only the vicinity of the endpoint $t' \approx t$. Substituting (24) into (21) and using the asymptotic form of the Green's function for $t' \to t$ (see Eq. (22)), we obtain:

$$\sum_{\nu} e^{-iE_\nu t}\varphi_\nu^{(n+1)}(\mathbf{r},t) = \iint G_{at}(\mathbf{r},t;\mathbf{r}',t')\times$$

$$\times V_{XUV}(\mathbf{r}',t')\sum_{\nu'} e^{-iE_{\nu'}t'}\varphi_{\nu'}^{(n)}(\mathbf{r}',t')d\mathbf{r}'dt'. \tag{26}$$

Next, approximating the slow functions $\varphi_\nu^{(n)}(\mathbf{r}',t')$ on the right-hand side of Eq. (26) by their values at $t' = t$ and using the relation between the stationary and time-dependent atomic Green's functions:

$$G_E(\mathbf{r},\mathbf{r}') = \int e^{iE(t-t')}G_{at}(\mathbf{r},t;\mathbf{r},t')dt', \tag{27}$$

we obtain:

$$\sum_{\nu} e^{-iE_\nu t}\varphi_\nu^{(n+1)}(\mathbf{r},t) =$$

$$= \sum_{\nu'} e^{-iE_{\nu'+1}t} G_{E_{\nu'+1}} V_+ \mid \varphi_{\nu'}^{(n)}\rangle +$$

$$+ \sum_{\nu'} e^{-iE_{\nu'-1}t} G_{E_{\nu'-1}} V_- \mid \varphi_{\nu'}^{(n)}\rangle. \tag{28}$$

Note that if the energy of the Green's function coincides with the energy of the ground state, then, according to perturbation theory, the Green's function is replaced by the reduced Green's function G'_{E_0} [69, 70]:

$$G'_{E_0} = \lim_{E\to E_0}\left[G_E(\mathbf{r},\mathbf{r}') - \frac{\varphi_0^*(\mathbf{r}')\varphi_0(\mathbf{r})}{E - E_0}\right]. \tag{29}$$

Given the weak dependence of the functions $\varphi_{\nu'}^{(n)}$ and $V_\pm$ on time [see (25)], we equate the coefficients of the "fast-oscillating" exponentials with identical exponents in (32) and obtain the equation for the functions $\varphi_\nu^{(n+1)}$:

$$\varphi_\nu^{(n+1)} = \begin{cases} G_{E_\nu} V_+ \mid \varphi_{\nu-1}^{(n)}\rangle + G_{E_\nu} V_- \mid \varphi_{\nu+1}^{(n)}\rangle, & \nu \neq 0, \\ G'_{E_0} V_+ \mid \varphi_{-1}^{(n)}\rangle + G'_{E_0} V_- \mid \varphi_{+1}^{(n)}\rangle, & \nu = 0. \end{cases} \tag{30}$$

The iterative method for solving Eq. (30) assumes the following expression for the zero iteration:

$$\varphi_\nu^{(0)}(\mathbf{r},t) = \varphi_0(\mathbf{r})\delta_{\nu,0}. \tag{31}$$

Thus, using the relations (24) and (30), one can find the nth-order correction for the slow part of the wave function, which formally coincides with the expression for the nth-order perturbation theory in a monochromatic field [70] (for a monochromatic field: $f_{XUV}(t) \equiv 1$, i.e., $V_\pm$ do not depend on time). We write the wave function $\Psi_n^{(s)}$ in the nth order of perturbation theory using the integral operator $\hat{P}_n$:

$$\Psi_n^{(s)}(\mathbf{r},t) = \hat{P}_n(E_0,t) \mid \varphi_0\rangle, \tag{32}$$

which represents the convolution of the atomic Green's function with all possible n-combinations of the operators V_+ and/or V_-. We present the explicit form of the operators $\hat{P}_n$ for the first three orders of perturbation theory ($n \leqslant 3$):

$$\hat{P}_0(E_0,t) = e^{-iE_0 t}\hat{I},$$

$$\hat{P}_1(E_0,t) = e^{-iE_1 t}G_{E_1}V_+ + e^{-iE_{-1}t}G_{E_{-1}}V_-,$$

$$\hat{P}_2(E_0,t) = e^{-iE_2 t}G_{E_2}V_+G_{E_1}V_+ +$$

$$+ e^{-iE_0 t}G'_{E_0}V_-G_{E_1}V_+ +$$

$$+e^{-iE_0t}G'_{E_0}V_+G_{E_{-1}}V_- +$$
$$+e^{-iE_{-2}t}G_{E_{-2}}V_-G_{E_{-1}}V_-,$$
$$\widehat{P}_3(E_0,t) = e^{-iE_3t}G_{E_3}V_+G_{E_2}V_+G_{E_1}V_+$$
$$+e^{-iE_1t}G_{E_1}V_-G_{E_2}V_+G_{E_1}V_+ +$$
$$+e^{-iE_1t}G_{E_1}V_+G'_{E_0}V_-G_{E_1}V_+ +$$
$$+e^{-iE_1t}G_{E_1}V_+G'_{E_0}V_+G_{E_{-1}}V_- +$$
$$+e^{-iE_{-1}t}G_{E_{-1}}V_-G'_{E_0}V_-G_{E_1}V_+ +$$
$$+e^{-iE_{-1}t}G_{E_1}V_-G'_{E_0}V_+G_{E_{-1}}V_- +$$
$$+e^{-iE_{-1}t}G_{E_{-1}}V_+G_{E_{-2}}V_-G_{E_{-1}}V_- +$$
$$+e^{-iE_{-3}t}G_{E_{-3}}V_-G_{E_{-2}}V_-G_{E_{-1}}V_-,$$

where $\hat{I}$ is the identity operator. It is evident that the slow part of the wave function is the sum of partial terms $\Psi_n^{(s)}(\mathbf{r},t)$:

$$\Psi^{(s)}(\mathbf{r},t) = \sum_{n=0}^{\infty}\Psi_n^{(s)}(\mathbf{r},t) = \sum_{n=0}^{\infty}\widehat{P}_n(E_0,t)\varphi_0(\mathbf{r}). \quad (33)$$

The similarity between the perturbation series in the XUV interaction of the atomic wave function in an intense IR field and the well-known perturbation result for the quasistationary atomic state in a monochromatic laser field, obtained within the quasistationary quasienergy state (QQES) method [70], is noteworthy. The series (33) formally coincides with the QQES result after replacing the exact quasienergy ε by the ground state energy E_0 and the field strength F_{XUV} by the instantaneous amplitude of the pulse field $\mathcal{F}_{XUV}(t) = F_{XUV}f_{XUV}(t)$. Thus, if the functional dependence of the QQES wave function $\Psi_{XUV}^{(QQES)}(\mathbf{r},t;\varepsilon,F_{XUV})$ on the quasienergy and field strength is known, the same dependence defines the wave function $\Psi^{(s)}(\mathbf{r},t)$:

$$\Psi^{(s)}(\mathbf{r},t) = \Psi_{XUV}^{(QQES)}(\mathbf{r},t;E_0,\mathcal{F}_{XUV}(t)). \quad (34)$$

In contrast to the slow part, the time dependence of the fast part $\Psi_n^{(r)}(\mathbf{r},t)$ is determined by the rapidly oscillating exponential factor $\sim e^{-iS}$, defined by the classical action S of the electron in the IR field along the closed trajectories (see Eqs. (5b), (9), and (11a)). It should be noted that, within the adiabatic approximation, the appearance of any products of two or more Volkov Green's functions exceeds the accuracy of the method. Therefore, in any order of perturbation theory for the XUV interaction, the expression for $\Psi_n^{(r)}(\mathbf{r},t)$ contains only one Volkov Green's function. From relation (21), it follows that $\Psi_n^{(r)}(\mathbf{r},t)$ is defined as the convolution result of either the slow part $\Psi_n(\mathbf{r},t)$ with the Green's function G_{IR}, approximated by the Volkov Green's function (see Eq. (22), or the fast part $\Psi_n(\mathbf{r},t)$ with the Green's function G_{IR}, approximated by the atomic Green's function (see Eq. (22)):

$$\Psi_{n+1}^{(r)}(\mathbf{r},t) =$$
$$= \int G_{vol}(0,t;0,t')\langle\psi_{\mathbf{K}'}^{(+)} \mid V_{XUV}(t') \mid \Psi_n^{(s)}(t')\rangle\psi_{\mathbf{K}}^{(+)}(\mathbf{r})dt' +$$
$$+ \iint G_{at}(\mathbf{r},t;\mathbf{r}',t')V_{XUV}(\mathbf{r}',t')\Psi_n^{(r)}(\mathbf{r}',t')d\mathbf{r}'dt', \quad (35)$$

where the first integral implies an approximate evaluation using the saddle-point method, while the second integral should be evaluated considering only the contribution from the vicinity of $t' \approx t$. Accordingly, $\Psi_{n+1}^{(r)}$ can be written as the sum of two terms:

$$\Psi_{n+1}^{(r)} = \Psi_{n+1}^{(r,1)} + \Psi_{n+1}^{(r,2)}, \quad (36)$$

where the expressions for для $\Psi_{n+1}^{(r,1)}$, $\Psi_{n+1}^{(r,2)}$ are discussed below.

We will use the approximate expression (27) for the Green's function $\Psi_n^{(s)}$ in the first integral of (39) and then perform the saddle-point integration over t'. As a result, we obtain an expression for $\Psi_{n+1}^{(r,1)}$:

$$\Psi_{n+1}^{(r,1)}(\mathbf{r},t) = \sum_{\nu}{}' \sum_{\tilde{t}'}\psi_{\tilde{\mathbf{K}}}^{(+)}(\mathbf{r})G_{\nu}(t,\tilde{t}') \times$$
$$\times\left[\langle\psi_{\tilde{\mathbf{K}}'}^{(+)} \mid V_+(\tilde{t}') \mid \varphi_{\nu-1}^{(n)}\rangle + \langle\psi_{\tilde{\mathbf{K}}'}^{(+)} \mid V_-(\tilde{t}') \mid \varphi_{\nu+1}^{(n)}\rangle\right], \quad (37)$$

where $\tilde{\mathbf{K}}' = \mathbf{K}'(t,\tilde{t}')$, $\tilde{\mathbf{K}} = \mathbf{K}(t,\tilde{t}')$, the summation is carried out over all allowed values of ν with the same parity as $n+1$ in the interval $|\nu| \leqslant n+1$, and the saddle points $\tilde{t}' \equiv \tilde{t}'(t)$, defined by the equation:

$$\tilde{\mathbf{K}}'^2 = 2E_{\nu}. \quad (38)$$

In (37), the following notation is used, defined as:

$$G_{\nu}(t,\tilde{t}') = -\frac{e^{-iS(t,\tilde{t}')-iE_{\nu}\tilde{t}'}}{2\pi(t-\tilde{t}')^{3/2}} \times$$
$$\times\left[\tilde{\mathbf{K}}' \cdot \mathbf{F}_{IR}(\tilde{t}') - 2E_{\nu} / (t-\tilde{t}')\right]^{-1/2}. \quad (39)$$

The accuracy of the approximate expression for the Green's function in (22) implies that only those

saddle points v should be considered in the sum (37) that ensure the solution of Eq. (38) in real numbers.

From expression (37), it follows that the correction $\Psi_{n+1}^{(r,1)}$ to the fast part of the wave function describes a rescattering state, formed within the three-step scenario:

First stage: as a result of the absorption or emission of n XUV photons, the atomic system forms states $\varphi_{\nu}^{(n)}$.

Second stage: a stimulated single-photon transition (with absorption or emission of a photon) from one of the states $\varphi_{\nu}^{(n)}$ to a continuum state with asymptotic momentum $\tilde{\mathbf{K}}'$ occurs at the moment $\tilde{t}'$.

Third stage: while in the continuum, the electron interacts with the intense IR field and forms, at the moment t, the state $\psi_{\tilde{\mathbf{K}}}^{(+)}(\mathbf{r})$, acquiring energy while moving along a closed classical trajectory. The propagation of the electron wave packet in the continuum is described by the multiplier $G_{\nu}(t,\tilde{t}')$.

As follows from Eqs. (5b) and (37), the fast part of the wave function is determined by the rapidly oscillating factor $e^{-iS(t,t')}$, which defines the IR-controlled propagation of the electron in the continuum, and the continuum state function $\psi_{\mathbf{K}}^{(+)}$. The same components determine the correction $\Psi_n^{(r,2)}$, so, without loss of generality, we represent $\Psi_n^{(r,2)}$ as:

$$\Psi_n^{(r,2)} = \sum_s \hat{a}_s^{(n)}(t) \mid \psi_{\mathbf{K}_s}^{(+)}\rangle, \tag{40}$$

where $\hat{a}_s^{(n)}(t) \propto e^{-iS(t,\hat{t}_s')}$ is a certain integral operator, and the summation is performed over all real ionization moments $\hat{t}_s'$, induced by the IR or XUV field. In the zero approximation for the XUV interaction: $\hat{t}_s' \equiv t_s'$ (see relation (7)), and $\hat{a}_s^{(0)} \equiv a_s(t)\hat{I}$ (see relation (9)). It is worth noting that the operator $\hat{a}_s^{(n)}(t)$ can be defined by two terms (denoted below by the indices a and b), describing two different scenarios of the electron interaction with the IR and XUV pulses:

a) The atomic electron tunnels into the IR-modified continuum and, while propagating along closed classical trajectories in the IR field, absorbs v and emits $n-v$ of XUV photons (parity of n and v is the same).

b) The atomic electron transitions into the IR-modified continuum with energy $E_{\nu'}$ by absorbing ν' of XUV photons, where, during propagation along the closed classical trajectories in the IR field, it absorbs and emits additional XUV photons.

According to the described mechanisms, we represent the function $\Psi_n^{(r,2)}$ as a sum:

$$\Psi_n^{(r,2)} = \Psi_n^{(r,2a)} + \Psi_n^{(r,2b)}. \tag{41}$$

The mathematical expression for the operator $\hat{a}_s^{(n)}(t)$, corresponding to the realization of Scenario (a), can be easily obtained from (35) (see the second integral term on the right-hand side), assuming that the "zero iteration" $\Psi_0^{(r)} = \Psi_{IR}^{(r)}$ for the fast part of the wave function is defined in (5). By sequentially calculating the time integrals in (35) and considering the contribution from the vicinity of the point $t' = t$, we obtain the general expression for $\Psi_n^{(r,2a)}(\mathbf{r},t)$:

$$\Psi_n^{(r,2a)}(\mathbf{r},t) = \sum_s a_s(t)\hat{P}_n\left(\frac{\mathbf{K}_s^2}{2},t\right)\psi_{\mathbf{K}_s}^{(+)}(\mathbf{r}), \tag{42}$$

where the summation includes all solutions of Eq. (7). It is worth noting that, similar to the previously considered case of the functions $\Psi^{(s)}(\mathbf{r},t)$, the summation of the perturbation series in n, taking into account the explicit form of (46), leads to a result formally coinciding with the expression for the quasienergy scattering state $\Psi_{\mathbf{K}}^{(+)}(\mathbf{r},t;\varepsilon,\mathcal{F}_{XUV}(t))$ of the atomic electron in a monochromatic XUV field [70] with the quasienergy $\varepsilon = \mathbf{K}_s^2/2$, the asymptotic momentum $\mathbf{K} = \mathbf{K}_s$, and the XUV field strength, equal to the instantaneous value $\mathcal{F}_{XUV}(t)$:

$$\sum_n \hat{P}_n\left(\frac{\mathbf{K}_s^2}{2},t\right)\psi_{\mathbf{K}_s}^{(+)}(\mathbf{r}) =$$

$$= \Psi_{\mathbf{K}_s}^{(+)}\left(\mathbf{r},t;\frac{\mathbf{K}_s^2}{2},\mathcal{F}_{XUV}(t)\right) \equiv \Psi_{\mathbf{K}_s}^{(+)}(\mathbf{r},t), \tag{43}$$

and therefore,

$$\Psi^{(r,2a)}(\mathbf{r},t) = \sum_n \Psi_n^{(r,2a)}(\mathbf{r},t) = \sum_s a_s(t)\Psi_{\mathbf{K}_s}^{(+)}(\mathbf{r},t). \tag{44}$$

The correction $\Psi_n^{(r,2b)}$ to the fast part of the wave function, responsible for the realization of Scenario (b), arises in the second and higher orders of perturbation theory. As the "zero iteration" for obtaining this correction, we use the term $\Psi_1^{(r,1)}$ ($\Psi_0^{(r)} = \Psi_1^{(r,1)}$), corresponding to the absorption of a XUV photon (see the first term in the square brackets in (37)):

$$\Psi_0^{(r)}(\mathbf{r},t) = \sum_{\tilde{t}'} \psi_{\tilde{\mathbf{K}}}^{(+)}(\mathbf{r}) G_1(t,\tilde{t}') \langle \psi_{\tilde{\mathbf{K}}'}^{(+)} \mid V_+(\tilde{t}') \mid \varphi_0 \rangle, \quad (45)$$

where $\tilde{t}'$ is determined from Eq. (38) with $\nu = 1$. Substituting (45) into the second term in (35) and evaluating the time integral in the vicinity of $t' \approx t$, we obtain the desired second-order correction:

$$\Psi_2^{(r,2b)}(\mathbf{r},t) = \sum_{\tilde{t}'} \hat{P}_1 \left(\frac{\tilde{\mathbf{K}}^2}{2}, t \right) \psi_{\tilde{\mathbf{K}}}^{(+)}(\mathbf{r}) \times$$

$$\times G_1(t,\tilde{t}') \langle \psi_{\tilde{\mathbf{K}}'}^{(+)} \mid V_+(\tilde{t}') \mid \varphi_0 \rangle. \quad (46)$$

It is easy to give a transparent physical meaning to relation (46): the electron, being in the bound state, absorbs a XUV photon and passes to a continuum state with asymptotic momentum $\tilde{\mathbf{K}}'$ (that corresponds to the matrix element $\langle \psi_{\tilde{\mathbf{K}}'}^{(+)} \mid V_+(\tilde{t}') \mid \varphi_0 \rangle$ in (46)). The electron propagates in the IR-dressed continuum along a closed trajectory (see the multiplier $G_1(t,\tilde{t}')$). As a result, it forms a continuum state at time t through a single-photon channel of interaction with XUV radiation (i.e., by absorbing or emitting a XUV photon).

In the third order of perturbation theory, the calculations are carried out similarly, and the corresponding correction takes the form:

$$\Psi_3^{(r,2b)}(\mathbf{r},t) = \sum_{\tilde{t}'(\nu=1)} \hat{P}_2 \left(\frac{\tilde{\mathbf{K}}^2}{2}, t \right) \psi_{\tilde{\mathbf{K}}}^{(+)}(\mathbf{r}) \times$$

$$\times G_1(t,\tilde{t}') \langle \psi_{\tilde{\mathbf{K}}'}^{(+)} \mid V_+(\tilde{t}') \mid \varphi_0 \rangle +$$

$$+ \sum_{\tilde{t}'(\nu=2)} \hat{P}_1 \left(\frac{\tilde{\mathbf{K}}^2}{2}, t \right) \psi_{\tilde{\mathbf{K}}}^{(+)}(\mathbf{r}) \times$$

$$\times G_2(t,\tilde{t}') \langle \psi_{\tilde{\mathbf{K}}'}^{(+)} \mid V_+(\tilde{t}') G_{E_1} V_+(\tilde{t}') \mid \varphi_0 \rangle, \quad (47)$$

where the times $\tilde{t}'$ for the first (second) sum are found from Eq. (38) for $v = 1$ ($v = 2$) respectively. The interpretation of the first sum in (47) is analogous to that provided for relation (46), except that at the final stage, the continuum state is formed through the two-photon interaction with the XUV radiation. The partial terms in the second sum reflect the following physical mechanism: the bound electron, having absorbed two photons, passes into a continuum state, where it propagates along a closed trajectory driven by the IR field and forms a continuum state through a single-photon channel of interaction with XUV radiation. It should be noted that, although the determination of higher-order corrections presents no significant difficulties, they are not considered in this work due to the complexity of the final expressions.

4. GENERATION OF RADIATION BY AN ATOM IN SYNCHRONIZED IR AND XUV PULSES

4.1. Generation channels

We will use the obtained expressions for the wave function to determine the radiation generation amplitude by an atom in the field of synchronized, linearly polarized IR and XUV pulses:

$$\mathbf{F}(t) = \mathbf{F}_{IR}(t) + \mathbf{F}_{XUV}(t - \tau), \quad (48)$$

where τ is the time delay between the pulses, defined as the time interval between the peaks of their envelopes. The amplitude of radiation generation is given by Eq. (12) with the substitution:

$$\Psi_{IR}(\mathbf{r},t) \to \Psi(\mathbf{r},t), \quad \tilde{\Psi}_{IR}(\mathbf{r},t) \to \tilde{\Psi}(\mathbf{r},t),$$

where $\tilde{\Psi}(\mathbf{r},t)$ is dual wave function, defined from $\Psi(\mathbf{r},t)$ by the same procedure as $\tilde{\Psi}_{IR}(\mathbf{r},t)$ (see discussion below Eq. (12)).

As shown in the previous section, the wave function $\Psi(\mathbf{r},t)$ is represented as the sum of "slow" ($\Psi^{(s)}(\mathbf{r},t)$) and "fast" ($\Psi^{(r)}(\mathbf{r},t)$) components. Accordingly, the radiation generation amplitude can be written as:

$$\mathcal{D}(\Omega) = \mathcal{D}^{(s)}(\Omega) + \mathcal{D}^{(r)}(\Omega) +$$

$$+ \tilde{\mathcal{D}}^{(r)}(\Omega) + \hat{\mathcal{D}}^{(r)}(\Omega), \quad (49a)$$

$$\mathcal{D}^{(s)}(\Omega) = \int \langle \tilde{\Psi}^{(s)}(\mathbf{r},t) \mid \mathbf{r} \mid \Psi^{(s)}(\mathbf{r},t) \rangle e^{i\Omega t} dt, \quad (49b)$$

$$\mathcal{D}^{(r)}(\Omega) = \int \langle \tilde{\Psi}^{(s)}(\mathbf{r},t) \mid \mathbf{r} \mid \Psi^{(r)}(\mathbf{r},t) \rangle e^{i\Omega t} dt, \quad (49c)$$

$$\tilde{\mathcal{D}}^{(r)}(\Omega) = \int \langle \tilde{\Psi}^{(r)}(\mathbf{r},t) \mid \mathbf{r} \mid \Psi^{(s)}(\mathbf{r},t) \rangle e^{i\Omega t} dt, \quad (49d$$

$$\hat{\mathcal{D}}^{(r)}(\Omega) = \int \langle \tilde{\Psi}^{(r)}(\mathbf{r},t) \mid \mathbf{r} \mid \Psi^{(r)}(\mathbf{r},t) \rangle e^{i\Omega t} dt, \quad (49e)$$

where each term is discussed in detail below.

The "slow" term $\mathcal{D}^{(s)}(\Omega)$ describes harmonic generation of the XUV field by the atomic system. Considering that $\Psi^{(s)}(\mathbf{r},t)$ is defined by the perturbation series (see Eqs. (33) and (34)), it is evident that $\mathcal{D}^{(s)}(\Omega)$ can be expressed in terms of nonlinear susceptibilities χ_n at the frequencies of the generated harmonics:

$$\mathcal{D}^{(s)}(\Omega) = e^{i\Omega\tau}\sum_n \chi_n(\omega_{XUV})F_{XUV}^n f_n(\Omega), \quad (50)$$

$$f_n(\Omega) = \frac{1}{2^n}\int_{-\infty}^{\infty} f_{XUV}^n(t)e^{i\Omega t}dt. \quad (51)$$

It is worth noting that, due to dipole selection rules for centrally symmetric systems, the nonlinear susceptibilities $\chi_n(\omega_{XUV})$ for even number n vanish. However, if we more accurately account for the IR-field effects in the zero-order approximation $\Psi_{IR}^{(0)}(\mathbf{r},t)$, it can be shown that the susceptibilities $\chi_n(\omega_{XUV})$ should be replaced by generalized nonlinear susceptibilities of the atomic system in a static electric field with a strength corresponding to the IR pulse at the delay time τ:

$$\chi_n(\omega_{XUV}) \to \chi_n^{(DC)}(\omega_{XUV}; \mathcal{F}_{DC} = F_{IR}(\tau)). \quad (52)$$

In this case, the prohibition on the generation of even harmonics is lifted, and the spectrum of the generated radiation exhibits peaks corresponding to the frequencies $N\omega_{XUV}$, where N is an integer (see, for example, [39]). Let us consider the "fast" term $\mathcal{D}^{(r)}(\Omega)$ in Eq. (49). Taking into account that the fast part of the wave function in the synchronized IR and XUV pulses is the sum of two terms (see Eq. (36)), we write $\mathcal{D}^{(r)}(\Omega)$ as:

$$\mathcal{D}^{(r)}(\Omega) = \mathcal{D}^{(r,1)}(\Omega) + \mathcal{D}^{(r,2)}(\Omega), \quad (53)$$

where $\mathcal{D}^{(r,i)}(\Omega)(i = 1,2)$ are determined by the corresponding corrections for the fast part of the wave function. Using Eqs. (32) and (37), we obtain $\mathcal{D}^{(r,1)}(\Omega)$ in the form:

$$\mathcal{D}^{(r,1)}(\Omega) = \int \mathcal{D}^{(r,1)}(t)e^{i\Omega t}dt, \quad (54a)$$

$$\mathcal{D}^{(r,1)}(t) = \sum_{\nu=1}^{\infty}\sum_{\tilde{t}'} M_\nu(\tilde{t}')G_\nu\langle\tilde{\Psi}^{(s)} \mid \mathbf{r} \mid \psi_{\tilde{\mathbf{K}}}^{(+)}\rangle, \quad (54b)$$

$$\mid \tilde{\Psi}^{(s)}\rangle = \sum_\nu [\hat{P}_\nu(E_0, -t)\varphi_0]^*, \quad (54c)$$

where $G_\nu \equiv G_\nu(t,\tilde{t}')$is defined in (39), $\tilde{t}'$ are roots of Eq. (38), and $M_\nu(\tilde{t}')$ is the sum of matrix elements describing the ν-photon XUV-induced excitation of the atomic system from the initial state φ_0 to the continuum state $\psi_{\tilde{\mathbf{K}}'}^{(+)}$, considering the re-emission channels. The explicit form of $M_\nu(\tilde{t}')$ can be determined within the perturbation theory using the recurrence relation (30):

$$M_1(\tilde{t}') = \langle\psi_{\tilde{\mathbf{K}}'}^{(+)} \mid \tilde{V}_+ \mid \varphi_0\rangle$$

$$+\langle\psi_{\tilde{\mathbf{K}}'}^{(+)} \mid \tilde{V}_+ G'_{E_0}\tilde{V}_- G_{E_0+\omega_{XUV}}\tilde{V}_+ \mid \varphi_0\rangle + \cdots,$$

$$M_2(\tilde{t}') = \langle\psi_{\tilde{\mathbf{K}}'}^{(+)} \mid \tilde{V}_+ G_{E_0+\omega_{XUV}}\tilde{V}_+ \mid \varphi_0\rangle + \cdots$$

$$M_3(\tilde{t}') = \langle\psi_{\tilde{\mathbf{K}}'}^{(+)} \mid \tilde{V}_+ G_{E_0+2\omega_{XUV}}\tilde{V}_+ G_{E_0+\omega_{XUV}}\tilde{V}_+ \mid \varphi_0\rangle + \cdots,$$

where $\tilde{V}_\pm \equiv V_\pm(\mathbf{r},\tilde{t}')$ (note that for the case of a linearly polarized XUV pulse, $\tilde{V}_+ = \tilde{V}_-$). The third factor in **(**54b**)** (the matrix element $\langle\tilde{\Psi}^{(s)} \mid \mathbf{r} \mid \psi_{\tilde{\mathbf{K}}}^{(+)}\rangle$) determines the amplitude of XUV-assisted recombination into the atomic state at the moment t (see expression **(**34**)**).

The function $\mathcal{D}^{(r,1)}(t)$ rapidly changes with variations in the time t due to the presence of the rapidly oscillating factor $e^{-iS(t,\tilde{t}')}$ in $G_\nu n$. Given that the time interval between ionization and recombination (i.e., the time of electron propagation in the continuum driven by the IR field) is on the order of the IR field period ($|\tilde{t} - \tilde{t}'| \sim T_{IR}$), ionization and recombination cannot occur throughout the duration $\mathcal{T}_{XUV}$ of the attosecond XUV pulse ($\mathcal{T}_{XUV} \ll T_{IR}$). This circumstance allows us to omit all terms in the sum over ν in (54c) except for $\nu = 0$, and to write the recombination amplitude $\langle\tilde{\Psi}^{(s)} \mid \mathbf{r} \mid \psi_{\tilde{\mathbf{K}}}^{(+)}\rangle$ in the lowest-order approximation in F_{XUV} (i.e., assuming $F_{XUV} = 0$ for the state $\tilde{\Psi}^{(s)}$):

$$\langle\tilde{\Psi}^{(s)} \mid \mathbf{r} \mid \psi_{\tilde{\mathbf{K}}}^{(+)}\rangle \approx e^{iE_0 t}\langle\varphi_0 \mid \mathbf{r} \mid \psi_{\tilde{\mathbf{K}}}^{(+)}\rangle. \quad (55)$$

Estimation of the integral (54a) by the stationary phase method leads to the result:

$$\mathcal{D}^{(r,1)}(\Omega) = \sum_{\tilde{t}} \mathcal{D}^{(r,1)}(\tilde{t})e^{i\Omega\tilde{t}}, \quad (56)$$

where the summation is performed over all times that satisfy the equation:

$$\frac{\tilde{\mathbf{K}}^2}{2} = \Omega + E_0, \quad \tilde{\mathbf{K}} = \mathbf{K}(\tilde{t},\tilde{t}'). \quad (57)$$

When solving this equation, one should take into account the implicit dependence of $\tilde{t}' = \tilde{t}'(\tilde{t})$ according to Eq. (38). Based on the obtained analytical relations, it is easy to give a physical interpretation of the radiation generation mechanism described by $\mathcal{D}^{(r,1)}(\Omega)$: the atomic electron absorbs ν XUV photons and passes into the continuum, where it propagates along a closed trajectory driven by of the intense IR field. At the moment of return to the atomic core, the energy gained by the electron is emitted as a photon with the frequency Ω through recombination into the ground state. This generation

mechanism is called the "XUV-initiated HHG channel" [44, 46, 71–73].

We represent the dipole moment $\mathcal{D}^{(r,2)}(\Omega)$, determined by the wave function $\Psi^{(r,2)}$, as the sum:

$$\mathcal{D}^{(r,2)}(\Omega) = \mathcal{D}^{(r,2a)}(\Omega) + \mathcal{D}^{(r,2b)}(\Omega), \quad (58)$$

where the partial dipole moments $\mathcal{D}^{(r,2a)}(\Omega)$ and $\mathcal{D}^{(r,2b)}(\Omega)$ correspond to the corrections $\Psi^{(r,2a)}$ and $\Psi^{(r,2b)}$ of the fast part of the wave function (see the discussion of Eq. (41)). Taking into account (42), we write $\mathcal{D}^{(r,2a)}(\Omega)$ as:

$$\mathcal{D}^{(r,2a)}(\Omega) = \int \mathcal{D}^{(r,2a)}(t) e^{i\Omega t}, \quad (59a)$$

$$\mathcal{D}^{(r,2a)}(t) = \sum_s a_s(t) \langle \tilde{\Psi}^{(s)} \mid \mathbf{r} \mid \Psi^{(+)}_{\mathbf{K}_s} \rangle, \quad (59b)$$

where $\Psi^{(+)}_{\mathbf{K}_s}$ is defined by relation (43). Considering the definition of the dual function $\tilde{\Psi}^{(s)}$, constructed from $\Psi^{(s)}$ (see relation (34)), we express the matrix element in (59b) as:

$$\langle \tilde{\Psi}^{(s)} \mid \mathbf{r} \mid \Psi^{(+)}_{\mathbf{K}_s} \rangle \approx \sum_n A^{(rec)}_n f^n_{XUV}(t) e^{-in\omega_{XUV} t}, \quad (60)$$

where $A^{(rec)}_n \propto F^n_{XUV}$ is the amplitude of photorecombination with the absorption $(n > 0)$ or emission $(n < 0)$ of $\mathbf{n}$ XUV photons. Since the function $a_s(t)$ is rapidly oscillating, the integral in (59a) can be evaluated using the stationary phase method. As a result, for the partial amplitude $\mathcal{D}^{(r,2a)}(\Omega)$, we obtain:

$$\mathcal{D}^{(r,2a)}(\Omega) = \sum_{n,s} a_s(t_s) A^{(rec)}_n \times$$
$$\times f^n_{XUV}(t_s - \tau) e^{i(\Omega - n\omega_{XUV})t_s}, \quad (61)$$

where the recombination times t_s are found from the stationary phase equation:

$$\frac{\mathbf{K}^2(t_s, t'(t_s))}{2} = \Omega + E_0 - n\omega_{XUV}, \quad (62)$$

and the corresponding ionization times $t'(t_s)$ satisfy Eq. (7) when substituting $t = t_s$. In the following, we will number possible solution pairs of the system of equations (7) and (62) with a single index s: (t_s, t'_s). The analytical relation (61) allows us to give a simple quasiclassical interpretation of the radiation generation mechanism described by the term $\mathcal{D}^{(r,2a)}(\Omega)$: at the moment t'_s, the bound electron tunnels and propagates along a closed trajectory until the moment of recombination t_s. Recombination occurs with the emission of a photon with frequency Ω, simultaneously with the absorption of $\mathbf{n}$ XUV photons. Moreover, the envelope of the XUV pulse acts as a "temporal separator", cutting off recombination moments for which the difference $|t_s - \tau|$ exceeds the duration of the XUV pulse. This radiation generation mechanism defines the XUV-assisted HHG channel [52, 53].

Now let us show that the remaining terms $\mathcal{D}^{(r,2b)}(\Omega)$, $\tilde{\mathcal{D}}^{(r)}(\Omega)$ and $\hat{\mathcal{D}}^{(r)}(\Omega)$ are negligibly small. The calculation of the partial dipole moment $\mathcal{D}^{(r,2b)}(\Omega)$, using relations (46), (47), and (33), shows that it is determined by terms that were discarded during the analysis of $\mathcal{D}^{(r,1)}(\Omega)$. In particular, the dipole matrix element of the transition between $\Psi^{(r,2b)}_2$ and $\Psi^{(s)}$ has a second order in F_{XUV} and defines a linear ($\propto F_{XUV}$) correction to the dipole moment in the one-photon XUV-initiated generation channel, through the XUV-interaction at the recombination step (i.e., it includes, along with the the XUV-initiated channel, also the one-photon XUV-assisted recombination channel). Similarly, it can be shown that $\Psi^{(r,2b)}_3$ gives a correction $\propto F^2_{XUV}$ to the one-photon XUV-initiated channel due to the two-photon interaction in the XUV-assisted channel, as well as a correction $\propto F_{XUV}$ to the two-photon XUV-initiated channel via the one-photon XUV-assisted mechanism. These corrections should be discarded due to the significant difference in the time scales between the dynamics of the atomic electron's interaction with the IR and attosecond XUV pulses: the characteristic time scale between sequential processes of ionization and recombination is comparable to the IR-field period. Therefore, the ionization and recombination stages cannot occur within the duration of a single attosecond XUV pulse.

To estimate the contribution of the dipole moment $\tilde{\mathcal{D}}^{(r)}(\Omega)$, defined by expression (49d), note that it describes the time-inverted process relative to the previously considered generation channels for the term $\mathcal{D}^{(r)}(\Omega)$. This directly follows from the definition of the dual wave function. For example, the generation of radiation in the XUV-assisted channel for $\tilde{\mathcal{D}}^{(r)}(\Omega)$ occurs under the following scenario: the bound electron emits radiation at the frequency Ω, with the simultaneous absorption of n XUV photons. As the result, the electron goes into a virtual state

with a larger negative energy and returns to the initial state, interacting with the intense IR field. Since all radiation formation stages occur at negative energy, within the quasiclassical approximation, this mechanism is strongly suppressed, and its contribution is negligibly small (see, for example, the discussion in [63]). Using similar reasoning for the XUV-initiated generation channel, we conclude that it can also be neglected.

Finally, the term $\hat{\mathcal{D}}^{(r)}(\Omega)$ must also be discarded in our consideration, as it is determined by the product of two fast parts of the wave function, and its inclusion exceeds the accuracy established in this analysis. Thus, we have shown that radiation generation by an atomic system, interacting with intense IR radiation and an attosecond XUV pulse, whose duration is much shorter than the IR field period, can occur within the framework of three channels: 1) XUV harmonic generation, defined by the corresponding atomic nonlinear susceptibilities; 2) the XUV-initiated generation channel; 3) the XUV-assisted generation channel.

4.2 Contribution of different radiation generation channels

Let us consider the general properties of the radiation generation channels, such as the position and width of the spectral region $[\Omega_{\min};\Omega_{\max}]$ for a given channel. These properties depend on the characteristics of the atomic target (the energy of the initial bound state) and the parameters of the laser field interacting with the atomic system. The contribution of different generation channels and their spectral overlap is of particular interest.

The frequency interval $[\Omega_{\min};\Omega_{\max}]$ can be determined from the requirement for the existence of real solutions to the saddle-point equations for the classical ionization and recombination times. To find them, we parametrize the electric field of the IR pulse through the vector potential $\mathbf{A}_{IR}(t)$:

$$\mathbf{F}_{IR}(t) = -\frac{\partial \mathbf{A}_{IR}(t)}{\partial t}, \tag{63a}$$

$$\mathbf{A}_{IR}(t) = -\mathbf{e}_x \frac{F_{IR}}{\omega_{IR}} f_{IR}(t)\sin(\omega_{IR}t), \tag{63b}$$

$$f_{IR}(t) = \begin{cases} \cos^2\left(\dfrac{\pi t}{\mathcal{T}_{IR}}\right) & |t| \le \dfrac{\mathcal{T}_{IR}}{2} \\ 0 & |t| \ge \dfrac{\mathcal{T}_{IR}}{2} \end{cases} \tag{63c}$$

where $\mathcal{T}_{IR} = 5T_{IR}$ is the pulse duration, $T_{IR} = 2\pi/\omega_{IR}$. In all numerical calculations, we assume the initial bound state energy $E_0 = -13.6$ eV, corresponding to the ground state of the hydrogen atom.

4.2.1 XUV-assisted channel

For the XUV-assisted generation channel, the ionization times $s_{t'}$ and recombination times s_t satisfy the system of equations (7) and (62):

$$\mathbf{K}'_s \cdot \dot{\mathbf{K}}'_s = 0, \quad \mathbf{K}^2(t'_s,t_s) = 2(\Omega + E_0 - n\omega_{XUV}). \tag{64}$$

As seen from Eq. (64), the solution of this system for an arbitrary n can be obtained from the solution for $n = 0$ by a corresponding frequency shift of the generated radiation: $\Omega \to \Omega + n\omega_{XUV}$. Therefore, below we analyze the case $n = 0$, which corresponds to harmonic generation in the absence of the XUV field. The system (64) has real solutions for $\Omega > |E_0|$ and $\Omega < \max \mathbf{K}^2/2 + |E_0| = \alpha_0 u_p + |E_0|$, where $u_p = F_{IR}^2/(4\omega_{IR}^2)$, α_0 is a numerical factor depending on the pulse envelope shape. For example, for a long monochromatic pulse ($f_{IR}(t) \equiv 1$), we obtain $\alpha_0 \approx 3.17$.

Fig. 1 shows the dependence of the frequency Ω of the generated radiation on the recombination times t_j. The color represents the absolute value of the tunneling factor $a_j^{(tun)}$, which enters the expression (14b) for the partial HHG amplitude for the IR field. It is seen from the figure that for fixed parameters of the laser pulse, the number of solutions of system (64) increases with decreasing Ω, which leads to the formation of a complex interference structure in the plateau region [17]. In the vicinity of the global maximum for Ω (i.e., the cutoff of the IR-induced HHG plateau), only two solutions exist, determining the well-known interference oscillations of the HHG yield near the cutoff region [74, 75].

4.2.2 XUV-initiated channel

For the XUV-initiated channel (consisting of an ν-photon transition of the electron from the ground state to the continuum, its laser-driven propagation, and subsequent recombination), the ionization and recombination times are determined by the following system of equations:

$$\frac{\tilde{\mathbf{K}}'^2}{2} = E_0 + \nu\omega_{XUV}, \tag{65a}$$

$$\frac{\tilde{\mathbf{K}}^2}{2} = E_0 + \Omega, \tag{65b}$$

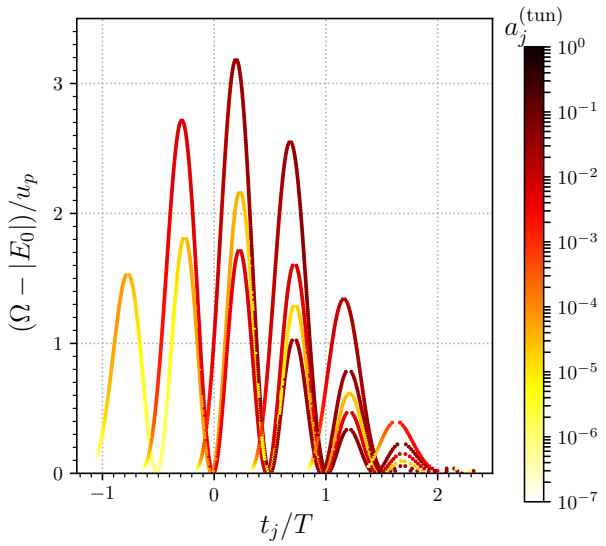

Fig. 1. Dependence of the recombination time on the frequency of the generated radiation for an IR pulse with a carrier frequency $\omega = 1$ eV and a peak intensity $2 \cdot 10^{14}$ W/cm^2. The color represents the value of the tunneling factor (10), calculated for the ionization and recombination times satisfying the system of equations (64) $u_p = 26.89$ eV. $|E_0| = 13.65$ eV.

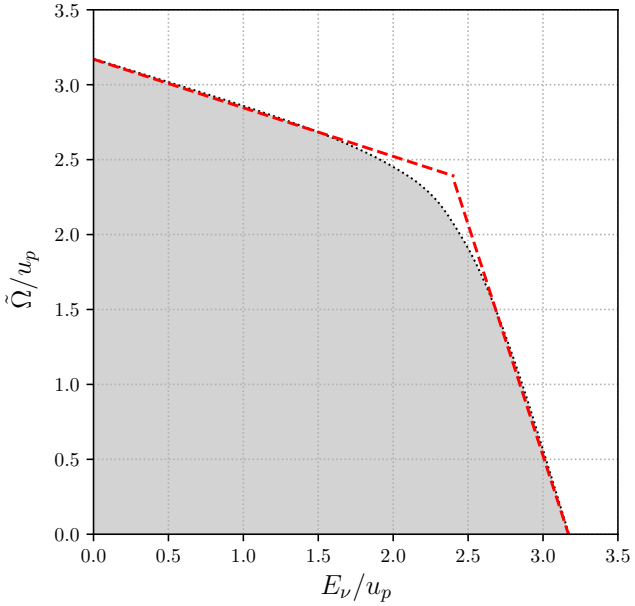

Fig. 2. Region of existence for solutions of the saddle-point equations (65). The dark gray area represents the parameter region obtained from the numerical solution, while the red dashed line shows the linear law (67).

where $\tilde{\mathbf{K}}' = \mathbf{K}'(\tilde{t},\tilde{t}')$, $\tilde{\mathbf{K}} = \mathbf{K}(\tilde{t},\tilde{t}')$.

Equation (65a) has real solutions under the following necessary condition:

$$\nu\omega_{XUV} \leqslant \max\left(|E_0| + \frac{\tilde{\mathbf{K}}'^2}{2}\right) = |E_0| + \alpha_0 u_p. \quad (66)$$

To determine the boundaries of the spectral region $[\Omega_{\min};\Omega_{\max}]$ of classically allowed frequencies of the generated radiation, note that the system of equations (65) is invariant with respect to the replacement $(\tilde{\mathbf{K}}', E_\nu) \leftrightarrow (\tilde{\mathbf{K}}, \tilde{\Omega})$, where $\tilde{\Omega} = \Omega + E_0$. The maximum values $\tilde{\mathbf{K}}'^2/2$ and $\tilde{\mathbf{K}}^2/2$ are identical due to the obvious symmetry in the dependence of $\mathbf{K}(t,t')$ and $\mathbf{K}'(t,t')$ on the times t, t' (see Eqs. (6) and (8)). Therefore, in the plane of the variables $\tilde{\Omega}$ and E_ν, the desired region of real solutions (or classically allowed energies $\tilde{\Omega}$ and E_ν is symmetric with respect to the line $\tilde{\Omega} = E_\nu$. Moreover, since the momenta $\mathbf{K}$ and $\mathbf{K}'$ are proportional to F_{IR}/ω_{IR}, the region of real solutions in the coordinates $\tilde{\Omega}$, E_ν scales by the magnitude u_p. From the above, it follows that the boundary of classically allowed energies can be expressed using a symmetric function $g(x,y) = g(y,x)$ of the two arguments $x = \tilde{\Omega}/u_p$ and $y = E_\nu/u_p$ in the form of the following equation:

$$g\left(\frac{\tilde{\Omega}}{u_p}, \frac{E_\nu}{u_p}\right) = 0.$$

Fig. 2 shows the region of classically allowed energies $\tilde{\Omega}$ and E_ν, obtained from the numerical analysis of the system of equations (65). The desired region is well approximated by two straight lines [76]:

$$g\left(\frac{\tilde{\Omega}}{u_p}, \frac{E_\nu}{u_p}\right) \approx \begin{cases} \dfrac{\tilde{\Omega}}{u_p} + \beta_0 \dfrac{E_\nu}{u_p} - \alpha_0, & \tilde{\Omega} > E_\nu, \\ \dfrac{E_\nu}{u_p} + \beta_0 \dfrac{\tilde{\Omega}}{u_p} - \alpha_0, & \tilde{\Omega} < E_\nu \end{cases} \quad (67)$$

where

$$\beta_0 = F_{IR}(t_0') / F_{IR}(t_0),$$

t_0' and t_0 are ionization and recombination times, corresponding to the global maximum of $\tilde{\mathbf{K}}^2/2$ (for the monochromatic field $\beta_0 = 0.324$).

The dependence of the solutions of the system of equations (65) on the frequency of the generated radiation Ω is shown in Fig. 3. Each pair of solutions $(\tilde{t}',\tilde{t})$ is represented by a point, the color of which corresponds to a specific value of ω_{XUV}. As can be seen from the presented figure, as well as from the above estimate (66), the number of real roots

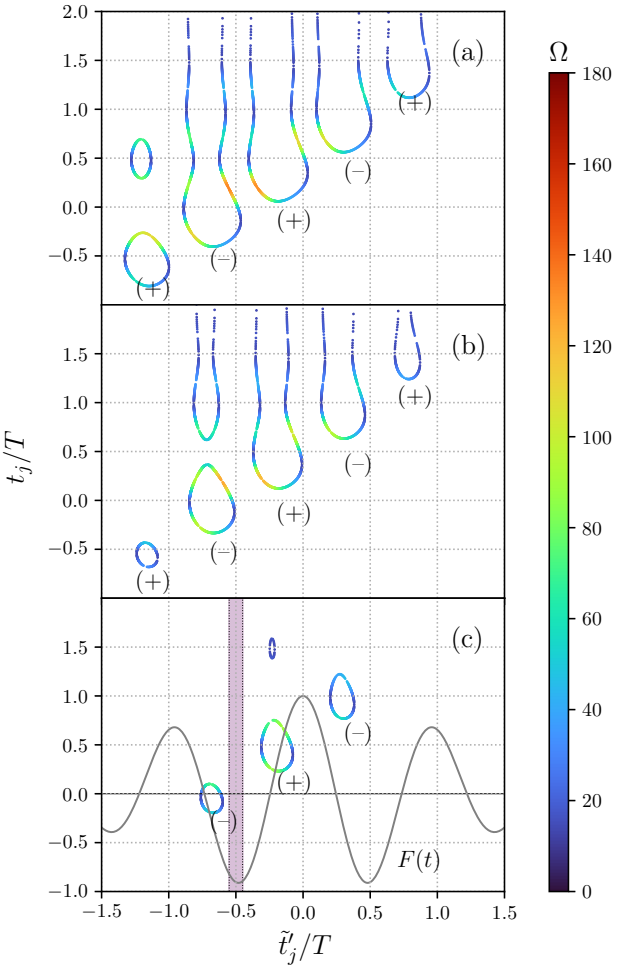

Fig. 3. Solutions of the system of equations (65) for the classical ionization times t'_+ and recombination times t_+ in the case of single-photon ionization ($\nu = 1$) and various values of the XUV photon energy: (a) $\omega_{XUV} = 40$ eV, (b) $\omega_{XUV} = 60$ eV, (c) $\omega_{XUV} = 100$ eV. The shaded purple area represents the duration of the XUV pulse. The initial-state energy, carrier frequency of the IR pulse, XUV pulse duration, and the time delay between the pulses are the same as in Fig. 1, while the peak intensity of the IR pulse is $I = 3 \cdot 10^{14}$ W/cm^2. The signs (±) indicate the direction of the instantaneous momentum $\tilde{\mathbf{K}}$ of the electron at the moment of ionization relative to the polarization vector $\mathbf{e}_x$ of the IR field: (+) for the case $(\tilde{\mathbf{K}} \cdot \mathbf{e}_x) > 0$ and (−) for the case $(\tilde{\mathbf{K}} \cdot \mathbf{e}_x) < 0$. The black line represents the profile of the IR field intensity in arbitrary units.

of the system (65) decreases with the increase of ω_{XUV}, while the region of possible values for the ionization and recombination times shrinks (see the regions bounded by closed curves in Fig. 3). We remind that in the theory being developed, the interaction of the XUV pulse with the atomic system is determined by the specific moments of ionization and recombination (see Section 3), which define the closed trajectory of the free electron in the IR field. Therefore, at certain time delays, the considered generation channel can be suppressed due to the lack of overlap between the solution region of the system (65) and the time interval of interaction with the XUV pulse. For example, in Fig. 3c, it is shown that for an XUV pulse with a time delay $\tau = -0.5T$, the region of acceptable values for $\tilde{t}'$ does not intersect with the time interval of the XUV pulse duration.

Dependence of the generated radiation frequency on the recombination times for all the discussed generation channels is shown in Fig. 3. The regions corresponding to elastic scattering of the XUV photon by the atomic system (Rayleigh scattering) and the second harmonic generation are indicated by the horizontal dashed lines in the figure (solid bold horizontal lines correspond to $\Omega = \omega_{XUV}$ and $\Omega = 2\omega_{XUV}$). At IR-pulse intensities of $I \gtrsim 2 \cdot 10^{14}$ W/cm^2, we observe a spectral overlap between the harmonic generation channel in the IR field (black bold lines in Fig. 3) and the elastic scattering channel of the XUV photon, leading to the specific oscillations in the harmonic generation spectrum [60]. As the IR pulse intensity increases, overlap with the XUV harmonic generation channels occurs (see Fig. 3(c), where overlap with the second XUV harmonic generation channel is observed at $I = 4 \cdot 10^{14}$ W/cm^2). We note, that for the occurrence of interference between different radiation generation channels, necessary conditions are spectral overlap of the channels and comparable generation probabilities within the desired channels. The XUV-initiated channels (green and orange lines) overlap spectrally only with the harmonic generation channel in the single IR pulse. Moreover, as seen in Fig. 3, as the number of photons in the XUV-initiated channel increases, the spectral overlap region shrinks, which is obviously related to the reduced energy gain by the electron during its propagation in the IR-field after absorbing ν XUV photons (see Fig. 2). Thus, the observation of XUV-initiated generation channels with $\nu > 1$ is difficult due to suppression by the more intense HHG channel in the absence of the XUV pulse and is possible only with a significant increase in the XUV-field intensity. For $\nu = 1$, the XUV-initiated channel can be distinguished under the orthogonal geometry of the IR and XUV pulses [77]. In contrast, the XUV-assisted channels (thin gray lines in Fig. 3) contribute to the generation of higher-frequency radiation, forming sequential plateau-like structures in the HHG spectra [53]. Typically, the XUV-assisted channels interfere with the harmonic generation channels of the XUV radiation

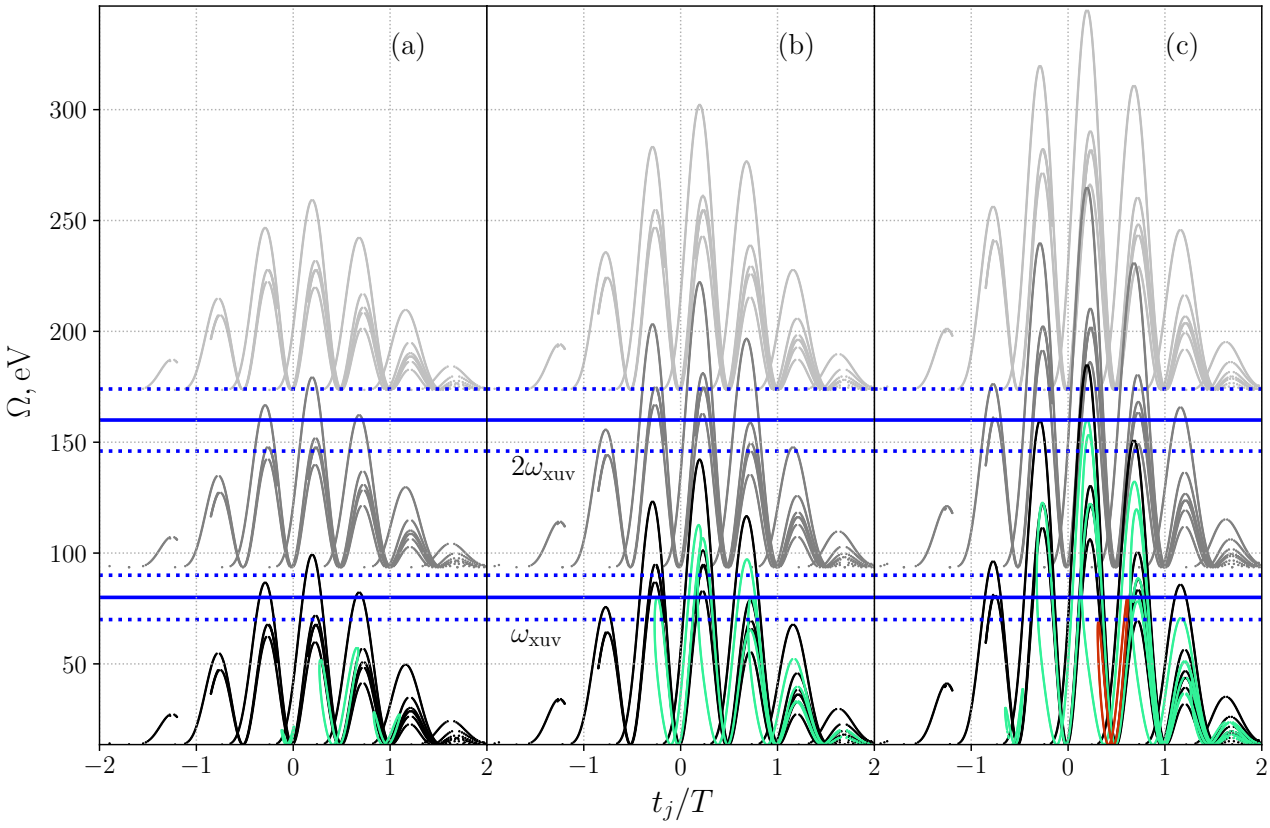

Fig. 4. Spectrograms of generation channels for different peak intensities of the IR pulse: (a) $I = 2 \cdot 10^{14}$ W/cm^2, (b) $3 \cdot 10^{14}$ W/cm^2 and (c) $4 \cdot 10^{14}$ W/cm^2. Black lines show the HHG channel in the absence of the XUV field. Gray lines represent the XUV-assisted channel with the absorption of one and two XUV photons. Green lines (orange lines) represent the XUV-initiated channel with the absorption of one (two) XUV photons. The solid horizontal lines indicate the values of $\Omega = N\omega_{XUV}$ ($N = 1,2$), while the dashed horizontal lines show the boundaries of the spectral regions for the first and second XUV harmonics. The XUV photon energy is $\omega_{XUV} = 80$ eV, and the initial-state energy, carrier frequency of the IR field, and XUV pulse envelope parameters are the same as in Fig. 1.

[39]. Note that it is possible to select conditions for spectral overlap between the harmonic generation channel in the single IR pulse and the XUV-assisted and XUV-initiated channels [see Fig. 4(c)].

5. CONCLUSION

In this work, an adiabatic approach has been developed for analyzing the interaction effects of an IR-dressed atomic system with a short (attosecond) XUV pulse. The nonlinear effects due to XUV interaction result in the emergence of additional XUV-induced radiation generation channels. Depending on the nature of the XUV pulse's influence on the atom, these channels can be classified into three types. The first one is the XUV-induced modification of nonlinear susceptibilities of the atomic system. For initially non-polarized targets, the odd-order susceptibilities (e.g., atomic polarizability, describing Rayleigh scattering of the XUV photon [60], or the third-order susceptibility, responsible for third XUV harmonic generation) do not vanish in the absence of the IR field. Thus, at moderate field intensities, they can be approximated by the susceptibilities of the free atom. In contrast, even-order susceptibilities (e.g., those describing the XUV rectification effect [61] or the second XUV harmonic generation [39]) vanish when the IR field is switched off, as they are caused by the IR-induced symmetry breaking of the atomic state. The first type of channels has been thoroughly studied in the works cited above.

The main focus of this study is on the second and third types of channels – the XUV-induced ionization channels, involving the absorption of XUV photons during the first stage of the three-step rescattering mechanism, and the XUV-assisted

recombination channels, where XUV photons are absorbed at the moment the electron returns to the atomic core. Analysis of the contributions of partial amplitudes associated with various closed classical trajectories of the electron in the IR field, in accordance with the described XUV-initiated and XUV-assisted channels, revealed that for moderate IR pulse intensities, the spectral region of XUV-initiated channels overlaps only with the harmonic spectrum of the IR field in the absence of the XUV pulse, while the probability of interference with XUV harmonic generation channels is negligibly small. In contrast, XUV photon absorption during recombination (XUV-assisted channel) significantly expands the spectrum of the generated radiation, enabling the interference of different generation channels.

It is worth noting that for short XUV pulses, the energy range of the generation channels strongly depends on the time delay between the XUV and IR pulses. For example, in the case of high XUV photon energies, there are delay intervals where the XUV-initiated generation channel is suppressed. It is important to emphasize that interference phenomena, caused by the spectral overlap of different XUV-induced generation channels with the IR-field HHG channel, are key to a deeper understanding of atomic photoprocesses occurring in the field of synchronized XUV and IR pulses. These phenomena can also form the basis for optical methods to extract the temporal profile of the IR pulse from the measured generation spectra [39, 77, 78].

FUNDING

This work was supported by the Russian Science Foundation (theory of radiation generation by atoms in IR and XUV pulses, Project No. 22-12-00223) and by a grant from the Ministry of Science and Higher Education of the Russian Federation (general perturbation theory in XUV-interaction for the IR-dressed atom, Project No. 075-15-2021-1351).

REFERENCES

1. *J. Caillat, J. Zanghellini, M. Kitzler, O. Koch, W. Kreuzer, and A. Scrinzi*, Phys. Rev. A **71**, 012712 (2005).
2. *D. Bauer and P. Koval*, Comp. Phys. Comm. **174**, 396 (2006).
3. *D. A. Telnov and S.-I. Chu*, Phys. Rev. A **80**, 043412 (2009).
4. *L. Greenman, P. J. Ho, S. Pabst, E. Kamarchik, D. Mazziotti, and R. Santra*, Phys. Rev. A **82**, 023406 (2010).
5. *D. A. Telnov, K. E. Sosnova, E. Rozenbaum, and S.-I. Chu*, Phys. Rev. A **87**, 053406 (2013).
6. *T. Sato and K. L. Ishikawa*, Phys. Rev. A, **88**, 023402 (2013).
7. *S. Patchkovskii and H. Muller*, Comp. Phys. Comm. **199**, 153 (2016).
8. *V. Tulsky and D. Bauer*, Comp. Phys. Comm. **251**, 107098 (2020).
9. *A. A. Romanov, A. A. Silaev, M. V. Frolov, and N. V. Vvedenskii*, Phys. Rev. A **101**, 013435 (2020).
10. *V. V. Strelkov*, Phys. Rev. A **74**, 013405 (2006).
11. *O. I. Tolstikhin, T. Morishita, and S. Watanabe*, Phys. Rev. A **81**, 033415 (2010).
12. *O. I. Tolstikhin and T. Morishita*, Phys. Rev. A **86**, 043417 (2012).
13. *Y. Okajima, O. I. Tolstikhin, and T. Morishita*, Phys. Rev. A **85**, 063406 (2012).
14. *M. V. Frolov, N. L. Manakov, A. A. Minina, A. A. Silaev, N. V. Vvedenskii, M. Y. Ivanov, and A. F. Starace*, Phys. Rev. A **99**, 053403 (2019).
15. *A. V. Flegel, N. L. Manakov, A. V. Sviridov, M. V. Frolov, L. Geng, and L.-Y. Peng*, Phys. Rev. A **102**, 063119 (2020).
16. *A. V. Sviridov, M. V. Frolov, S. V. Popruzhenko, L. Geng, and L.-Y. Peng*, Phys. Rev. A **106**, 033117 (2022).
17. *A. V. Flegel, N. L. Manakov, I. V. Breev, and M. V. Frolov*, Phys. Rev. A **104**, 033109 (2021).
18. *A. A. Romanov, A. A. Silaev, T. S. Sarantseva, M. V. Frolov, and N. V. Vvedenskii*, New J. Phys. **23**, 043014 (2021).
19. *D. B. Milosevic and F. Ehlotzky*, Adv. At., Mol. Opt. Phys., **49**, 373 (2003).
20. *W. Becker, F. Grasbon, R. Kopold, D. B. Milosevic, G. G. Paulus, and H. Walther*, Adv. At. Mol. Opt. Phys. **48**, 35 (2002).
21. *A. Galstyan, O. Chuluunbaatar, A. Hamido, Y. V. Popov, F. Mota-Furtado, P. F. O'Mahony, N. Janssens, F. Catoire, and B. Piraux*, Phys. Rev. A **93**, 023422 (2016).
22. *Y. Popov, A. Galstyan, F. Mota-Furtado, P. F. O'Mahony, and B. Piraux*, ' Eur. Phys. J. D **71**, 93 (2017).
23. *L. V. Keldysh*, JETP **20**, 1307 (1965).
24. *M. V. Fedoruk*, Saddle point method, Nauka, Moscow (1977) [in Russian].
25. *P. Salieres, B. Carre, L. Le Deroff, F. Grasbon, G. G. Paulus, H. Walther, R. Kopold, W. Becker, D. B. Milosevic, A. Sanpera, and M. Lewenstein*, Science **292**, 902 (2001).
26. *D. B. Milosevic*, Phys. Rev. A **96**, 023413 (2017).
27. *P. B. Corkum*, Phys. Rev. Lett. **71**, 1994 (1993).

28. *W. Becker, A. Lohr, and M. Kleber*, J. Phys. B: At. Mol. Opt. Phys. **27**, L325 (1994).

29. *M. V. Frolov, N. L. Manakov, T. S. Sarantseva, M. Y. Emelin, M. Y. Ryabikin, and A. F. Starace*, Phys. Rev. Lett. **102**, 243901 (2009).

30. *M. V. Frolov, N. L. Manakov, and A. F. Starace*, Phys. Rev. A **82**, 023424 (2010).

31. *A. D. Shiner, B. E. Schmidt, C. Trallero-Herrero, H. J. Wörner, S. Patchkovskii, P. B. Corkum, J.-C. Kieffer, F. Legare, and D. M. Villeneuve*, Nat. Phys. **7**, 464 (2011).

32. *A. D. Shiner, B. E. Schmidt, C. Trallero-Herrero, P. B. Corkum, J.-C. Kieffer, F. L'egar'e, and D. M. Villeneuve*, J. Phys. B: At. Mol. Opt. Phys. **45**, 74010 (2012).

33. *V. N. Ostrovsky and D. A. Telnov*, J. Phys. B: At. Mol. Opt. Phys. **20**, 2397 (1987).

34. *V. N. Ostrovsky and D. A. Telnov*, J. Phys. B: At. Mol. Opt. Phys. **20**, 2421 (1987).

35. *M. Pont, R. Shakeshaft, and R. M. Potvliege*, Phys. Rev. A **42**, 6969 (1990).

36. *D. A. Telnov*, J. Phys. B: At. Mol. Opt. Phys. **24**, 2967 (1991).

37. *M. Pont, R. M. Potvliege, R. Shakeshaft, and Z. Teng*, Phys. Rev. A **45**, 8235 (1992).

38. *A. V. Flegel and M. V. Frolov*, J. Phys. A: Math. Theor. **56**, 505304 (2023).

39. *A. A. Romanov, A. A. Silaev, N. V. Vvedenskii, I. V. Breev, A. V. Flegel, and M. V. Frolov*, Phys. Rev. A **106**, 063101 (2022).

40. *K. Ishikawa*, Phys. Rev. Lett. **91**, 043002 (2003).

41. *K. L. Ishikawa*, Phys. Rev. A **70**, 013412 (2004).

42. *K. Schiessl, E. Persson, A. Scrinzi, and J. Burgdörfer*, Phys. Rev. A **74**, 053412 (2006).

43. *S. V. Popruzhenko, D. F. Zaretsky, and W. Becker*, Phys. Rev. A **81**, 063417 (2010).

44. *K. J. Schafer, M. B. Gaarde, A. Heinrich, J. Biegert, and U. Keller*, Phys. Rev. Lett. **92**, 023003 (2004).

45. *M. B. Gaarde, K. J. Schafer, A. Heinrich, J. Biegert, and U. Keller*, Phys. Rev. A 72, 013411(2005).

46. *J. Biegert, A. Heinrich, C. P. Hauri, W. Kornelis, P. Schlup, M. P. Anscombe, M. B. Gaarde, K. J. Schafer, and U. Keller*, J. Mod. Opt. **53**, 87 (2006).

47. *C. Figueira de Morisson Faria, P. Salieres, P. Villain, and M. Lewenstein*, Phys. Rev. A **74**, 053416 (2006).

48. *G.-T. Zhang, J. Wu, C.-L. Xia, and X.-S. Liu*, Phys. Rev. A **80**, 055404 (2009).

49. *M. R. Miller, C. Hernandez-Garcia, A. Jaron-Becker, and A. Becker*, Phys. Rev. A **90**, 053409 (2014).

50. *P. B. Corkum, N. H. Burnett, and M. Y. Ivanov*, Opt. Lett. **19**, 1870 (1994).

51. A. Fleischer and N. Moiseyev, Phys. Rev. A 77, 010102 (2008).

52. *A. Fleischer*, Phys. Rev. A **78**, 053413 (2008).

53. *T. S. Sarantseva, M. V. Frolov, N. L. Manakov, A. A. Silaev, N. V. Vvedenskii, and A. F. Starace*, Phys. Rev. A **98**, 063433 (2018).

54. *C. Buth, F. He, J. Ullrich, C. H. Keitel, and K. Z. Hatsagortsyan*, Phys. Rev. A **88**, 033848 (2013).

55. *A. C. Brown and H. W. van der Hart*, Phys. Rev. Lett. **117**, 093201 (2016).

56. *J.-A. You, J. M. Dahlstr..om, and N. Rohringer*, Phys. Rev. A **95**, 023409 (2017).

57. *J. Leeuwenburgh, B. Cooper, V. Averbukh, J. P. Marangos, and M. Ivanov*, Phys. Rev. Lett. **111**, 123002 (2013).

58. *J. Leeuwenburgh, B. Cooper, V. Averbukh, J. P. Marangos, and M. Ivanov*, Phys. Rev. A **90**, 033426 (2014).

59. *C. Buth, M. C. Kohler, J. Ullrich, and C. H. Keitel*, Opt. Lett. **36**, 3530 (2011).

60. *A. A. Romanov, A. A. Silaev, N. V. Vvedenskii, A. V. Flegel, and M. V. Frolov*, Opt. Lett. **47**, 3147 (2022).

61. *A. V. Flegel and M. V. Frolov*, Phys. Rev. Lett. **131**, 243202 (2023).

62. *A. A. Minina, M. V. Frolov, A. N. Zheltukhin, and N. V. Vvedenskii*, Quant. Elect. **47**, 216 (2017).

63. *M. Y. Kuchiev and V. N. Ostrovsky*, Phys. Rev. A **60**, 3111 (1999).

64. *M. V. Frolov, A. V. Flegel, N. L. Manakov, and A. F. Starace*, Phys. Rev. A **75**, 063407 (2007).

65. *R. M. Potvliege and R. Shakeshaft*, Phys. Rev. A **40**, 3061 (1989).

66. *N. L. Manakov, M. V. Frolov, A. F. Starace, and I. I. Fabrikant*, J. Phys. B: At. Mol. Opt. Phys. **33**, R141 (2000).

67. *N. L. Manakov and A. G. Fainshtein*, Theoret. and Math. Phys. **48**, 815 (1981).

68. *E. A. Pronin, A. F. Starace, M. V. Frolov, and N. L. Manakov*, Phys. Rev. A **80**, 063403 (2009).

69. *P. W. Langhoff, S. T. Epstein, and M. Karplus*, Rev. Mod. Phys. **44**, 602 (1972).

70. *N. L. Manakov, V. D. Ovsiannikov, and L. P. Rapoport*, Phys. Rep. **141**, 320 (1986).

71. *G. Gademann, F. Kelkensberg, W. K. Siu, P. Johnsson, M. B. Gaarde, K. J. Schafer, and M. J. J. Vrakking*, New J. Phys. **13**, 033002 (2011).

72. *D. Azoury, M. KrGjger, G. Orenstein, H. R. Larsson, S. Bauch, B. D. Bruner, and N. Dudovich*, Nat. Comm. **8**, 1453 (2017).

73. *M. Krüger, D. Azoury, B. D. Bruner, and N. Dudovich*, Appl. Sci. **9**, 378 (2019).

74. *M. V. Frolov, N. L. Manakov, T. S. Sarantseva, and A. F. Starace*, J. Phys. B: At. Mol. Opt. Phys. **42**, 035601 (2009).

75. *M. V. Frolov, N. L. Manakov, A. M. Popov, O. V. Tikhonova, E. A. Volkova, A. A. Silaev,*

N. V. Vvedenskii, and A. F. Starace, Phys. Rev. A **85**, 033416 (2012).

76. *A. V. Flegel, M. V. Frolov, A. N. Zheltukhin, and N. V. Vvedenskii*, Quant. Electron. **47**, 222 (2017).

77. *T. S. Sarantseva, A. A. Romanov, A. A. Silaev, N. V. Vvedenskii, and M. V. Frolov*, Phys. Rev. A **107**, 023113 (2023).

78. *A. A. Romanov, A. A. Silaev, T. S. Sarantseva, A. V. Flegel, N. V. Vvedenskii, and M. V. Frolov*, Opt. Lett. **48**, 3583 (2023).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 42–45

NUCLEI, PARTICLES, FIELDS, GRAVITATION, AND ASTROPHYSICS

CORRECTIONS FROM NON-LOCAL GRAVITY FOR BLACK HOLE SHADOW IMAGES

© 2025 S. O. Alexeyev[a, c], A. V. Nemtinova[b], O. I. Zenin[c, *], A. A. Baiderin[c]

[a]Sternberg Astronomical Institute, Lomonosov Moscow State University, Moscow, Russia
[b]Ural Federal University named after the First President of Russia B.N. Eltsin, Yekaterinburg, Russia
[c]Quantum Theory and High Energy Physics Department, Physics Faculty, Lomonosov Moscow State University, Moscow, Russia
*e-mail: dkiiiabu4@gmail.com

Received August 22, 2023

Revised September 13, 2024

Accepted September 14, 2024

Abstract. The work presents the necessary conditions for the "bounce" of the scale factor existence (except Big Bang at the initial moment of the Universe). We study rather wide range of parameter values. This fact seems to be significant both for the further construction of the theory of quantum gravity and for the consideration of subsequent cosmological evolution based on this model.

Keywords: *bounce, Fab Four, Horndeski model, scale factor, quantum gravity, scalar field*

DOI: 10.31857/S00444510250103e2

1. INTRODUCTION

Currently, General Relativity (GR) accurately explains nearly the entire body of astronomical data. However, from the very first cosmological solutions [1], Einstein's equations must necessarily include the energy-momentum tensor on the right-hand side. One approach is that the entire array of modern astrophysical data is well described by GR equations, and theories of gravity extending GR in various ways [2, 3, 4, 5, 6] are developed specifically to explain the physical nature of the right-hand side and its source.

One promising direction for extending GR has been scalar-tensor theories of gravity, where, as the name suggests, physical fields are included alongside geometric terms and curvature invariants. To address the issue of higher-order differential equations, theories have been constructed where higher degrees mutually cancel out, with the most general example of this approach being the Horndeski model [7, 8]. Despite significant constraints on the Horndeski model from gravitational-wave astronomy data [9, 10], interest in it (and theories derived from it that pass the GW170817 test) remains strong. This model has also been used to create nonsingular cosmology models, where the initial singularity is replaced by a "bounce" of the scale factor [11, 12]. This approach appears promising, and within the Horndeski framework, models known as the "Fab Four" were proposed, where the corrections themselves, without additional tuning parameters like the cosmological constant (Λ), ensure the accelerated expansion of the Universe [13, 14]. Nonsingular cosmological solutions within the Fab Four model, as an example of a scalar-tensor theory with a simpler structure than the general Horndeski theory, have also been discussed earlier [15].

The idea of adding quantum-field corrections to gravity models [16] allows, for example, the limitation of nonlocality size in gravity theories at the quantum limit [17]. This approach was also applied to the Fab Four model [18], and the additional inclusion of quantum-field corrections ensures that the speed of gravitational wave propagation now matches the experimental results of gravitational-wave astronomy. All of this highlights the potential of scalar-tensor models.

Therefore, we consider a nonminimal effective model of scalar-tensor gravity with third- and fourth-order field terms, formed by summing one-loop interactions [19] in the form:

$$S = \int \sqrt{-g}\left[\left(\frac{2}{\kappa^2} + \alpha\phi^2\right)R + \kappa^2\beta G^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \right.$$
$$\left. -\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{3!}\lambda\phi^3 - \frac{1}{4!}g\phi^4\right]d^4x, \qquad (1)$$

where $\kappa^2 = 32\pi G$, G is the Newtonian constant, ϕ is the new scalar field, R is the scalar curvature, α and β are dimensionless constants, λ is the cubic scalar coupling with mass dimension, g is the dimensionless fourth-order scalar coupling, and $G_{\mu\nu}$ is the Einstein tensor $\left(G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right)$. Despite its "extended" nature, this model remains significantly simpler than the standard version of the Horndeski or DHOST theory, increasing interest in its potential to explain dark energy and early Universe processes. To further analyze the applicability of this model to early Universe evolution, it is necessary to study its predictions for bounce and genesis realization [20]. This paper is dedicated to the first step in this direction – investigating the conditions for bounce existence. It is important to note that the absence of an initial singularity in the cosmological model significantly increases its appeal. For example, consider the search for parameter spaces where a "bounce" occurs [21] in second-order curvature correction gravity – the Gauss-Bonnet model [22, 23], one of the candidates for the semiclassical limit of string gravity [24]. Moreover, the bounce already appears with the simple addition of a scalar field, as in the Brans-Dicke model [25]. Thus, the presence of a nonsingular asymptotic solution in the considered theory serves as an additional argument for its relevance. As the first step in examining the strengths and weaknesses of the theory (1), we investigate this issue. Since additional constraints on the theory's parameters were previously proposed to pass astronomical tests (discussed at the end of Section 3) [19], it is of interest to compare these constraints with those imposed by the bounce requirement.

This paper is structured as follows. Section 2 derives the field equations for the theory proposed in [19]; Section 3 explores the parameter space constraints imposed by the bounce requirement; and Section 4 discusses the results and conclusions.

2. FIELD EQUATIONS

The Klein-Gordon equations are obtained by varying the action (1) with respect to the scalar field. Following [26], we have:

$$-\frac{1}{2!}\lambda\phi^2 - \frac{1}{3!}g\phi^3 + \Box\phi +$$
$$+2\alpha\phi R - 2\kappa^2\beta G^{\mu\nu}\nabla_\mu\nabla_\nu\phi = 0. \qquad (2)$$

Varying with respect to the metric tensor and introducing the effective gravitational constant $G_{eff}(\phi)$, which depends only on the scalar field, gives:

$$\frac{2}{\kappa^2} + \alpha\phi^2 = \frac{1}{16\pi G_{eff}(\phi)}. \qquad (3)$$

As a result, Einstein's equation takes the form:

$$\mathcal{G}_{\mu\nu} = \frac{1}{16\pi G_{eff}}G_{\mu\nu} - \left(\nabla_{\mu\nu} - g_{\mu\nu}\Box\right)\alpha\phi^2 - \frac{1}{2}\nabla_\mu\phi\nabla_\nu\phi -$$
$$-\frac{1}{2}g_{\mu\nu}\left(\frac{1}{2}(\nabla\phi)^2 + \frac{1}{3!}\lambda\phi^3 + \frac{1}{4!}g\phi^4\right) -$$
$$-\kappa^2\beta\left(-\nabla_\lambda\nabla_\mu\phi\nabla^\lambda\nabla_\nu\phi + \nabla_\mu\nabla_\nu\phi\Box\phi - \right.$$
$$-R_{\alpha\mu\nu\beta}\nabla^\alpha\phi\nabla^\beta\phi - \frac{1}{2}\left[\nabla_\mu\phi G_{\nu\lambda}\nabla^\lambda\phi + \right.$$
$$\left.+\nabla_\nu\phi G_{\mu\lambda}\nabla^\lambda\phi\right] - \frac{1}{2}\left[\nabla_\mu\phi R_{\nu\lambda}\nabla^\lambda\phi + \right.$$
$$\left.+\nabla_\nu\phi R_{\mu\lambda}\nabla^\lambda\phi\right] + g_{\mu\nu}\left[R^{\alpha\beta}\nabla_\alpha\phi\nabla_\beta\phi - \right.$$
$$\left.\left.\left.-\frac{1}{2}(\Box\phi)^2 + \frac{1}{2}(\nabla_{\alpha\beta}\phi)^2\right]\right]\right) = \frac{1}{2}T_{\mu\nu}, \qquad (4)$$

where $T_{\mu\nu}$ is the effective energy-momentum tensor:

$$T_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta(\sqrt{-g}L_m)}{\delta g^{\mu\nu}}, \qquad (5)$$

Here L_m is the matter Lagrangian.

3. COSMOLOGICAL SOLUTION WITH A "BOUNCE"

Following [22, 23], we consider an isotropic (Friedmann-like) cosmological solution of the form:

$$ds^2 = dt^2 - a^2(t)(dx^2 + dy^2 + dz^2), \qquad (6)$$

where both the scale factor a, and the scalar field ϕ depend only on the time coordinate t.

To study the behavior at the bounce point, we examine the system (2)–(4). At the bounce point, the scale factor must be positive and finite, i.e., $a = const > 0$. To ensure the scale factor reaches a minimum at the bounce point and to avoid a cosmological singularity $a = 0$ at any other point, it is necessary that $\dot{a} = 0$ and $\ddot{a} > 0$. With this, Einstein's equations at the bounce point can be rewritten as:

$$\frac{3}{4}\dot{\phi}^2 = -\frac{1}{12}\lambda\phi^3 - \frac{1}{48}g\phi^4, \tag{7}$$

$$-2\frac{\ddot{a}}{a}\left(\frac{2}{\kappa^2} + \alpha\phi^2\right) - 2\alpha\phi\ddot{\phi} + \frac{1}{4}\dot{\phi}^2 - \\ -5\kappa^2\beta\frac{\ddot{a}}{a}\dot{\phi}^2 + \frac{1}{12}\lambda\phi^3 + \frac{1}{48}g\phi^4 = 0, \tag{8}$$

The Klein-Gordon-Fock equation (2) takes the form:

$$\ddot{a} = \frac{a}{12\alpha\phi}\left(\ddot{\phi} - \frac{1}{2}\lambda\phi^2 - \frac{1}{6}g\phi^3\right). \tag{9}$$

If we consider the case where the energy-momentum tensor is represented by the scalar field, its absence would imply the absence of a nontrivial cosmological solution: $\phi = 0 \Rightarrow a = 0$. Since this would lead to the singularity we aim to avoid, we introduce the additional conditions:

$\dot{\phi} = 0$, $\phi = const > 0$ and $\ddot{\phi} > 0$.

From equation (8) and (9), we obtain an equation for the scalar field:

$$\phi = -4\frac{\lambda}{g}.$$

From equations (8) and (9), we derive an expression for the second derivative of the scalar field:

$$\ddot{\phi} = -\frac{\lambda}{36\alpha}\left(\frac{\frac{1}{\alpha\kappa^2} + 8\frac{\lambda^2}{g^2}}{1 + \frac{1}{12\alpha} + \frac{g^2}{96\kappa^2\alpha^2\lambda^2}}\right).$$

The final system of inequalities (after substituting into (9) with equations (7) and (8)) is:

$$\phi = -4\frac{\lambda}{g} > 0, \tag{10}$$

$$\ddot{\phi} = -\frac{\lambda}{36\alpha}\left(\frac{\frac{1}{\alpha\kappa^2} + 8\frac{\lambda^2}{g^2}}{1 + \frac{1}{12\alpha} + \frac{g^2}{96\kappa^2\alpha^2\lambda^2}}\right) > 0, \tag{11}$$

$$a > 0, \tag{12}$$

$$\ddot{a} = \frac{ag}{1728\alpha^2}\left(\frac{\frac{1}{\alpha\kappa^2} + 8\frac{\lambda^2}{g^2}}{1 + \frac{1}{12\alpha} + \frac{g^2}{96\kappa^2\alpha^2\lambda^2}}\right) + \frac{a\lambda^2}{18\alpha g} > 0. \tag{13}$$

From inequality (10), we obtain that λ and g must have opposite signs. It is also necessary for the stability of the model that $g > 0$. Otherwise, the scalar potential would be unbounded from below, rendering the model unstable. From inequality (11), it follows that $\lambda < 0$, then $\alpha > 0$. The final inequality (13) is automatically satisfied under conditions (10)–(12). We can also consider the case $\alpha < 0$. From (13), we obtain:

$$\left(\frac{\frac{1}{\alpha\kappa^2} + 8\frac{\lambda^2}{g^2}}{1 + \frac{1}{12\alpha} + \frac{g^2}{96\kappa^2\alpha^2\lambda^2}}\right) > -\frac{1728\alpha\lambda^2}{18g^2} > 0.$$

This implies that the expression inside the parentheses is positive. Thus, condition (11) also holds if $\lambda > 0$ and $g < 0$. However, this condition contradicts the necessary stability condition of the model. Therefore, these conditions are not suitable for the given problem.

4. CONCLUSION AND FINDINGS

In the non-minimal effective model of scalar-tensor gravity with third- and fourth-order field terms formed by summing one-loop interactions [19], the realization of a "bounce" solution is possible. The necessary conditions for the realization of the bounce solution are as follows: parameters $\lambda < 0$, $g > 0$ and $\alpha > 0$. A similar model was previously studied in [27], where $\alpha = 0$, the scalar field ϕ was absent, but the cosmological constant Λ was present, ensuring the same effect. The bounce solution is realized under the conditions $\Lambda = 0$ (although the case when $\lambda = g = 0$ is not possible in our model), $\rho = 0$ (similarly, in our case, the volume density is zero), $a_0 > 0$ (in our case, the scale factor $a > 0$) and $\beta < 0$ (which does not contradict our conditions). Thus, our results partially coincide with those previously obtained for a simpler version of the discussed model, except for the zero value of the cosmological constant and the parameter α (which was initially zero in the simpler version of the theory).

Thus, in the discussed scalar-tensor gravity model, instead of an initial singularity, a bounce is possible even in the simplest configuration, provided the initial constraints are met. This means that the model, with a simpler structure than most scalar-tensor models based on Horndeski's theory, not only solves the initial singularity problem but also brings us closer to the development of quantum gravity while offering the potential for the realization of both bounce and genesis scenarios.

FUNDING

The work of O.I.Z. was funded by the Foundation for the Advancement of Theoretical Physics and Mathematics "BASIS," grant 22-2-2-11-1.

REFERENCES

1. *A. Friedmann*, *Über die Krümmung des Raumes (About the curvature of space)*, Z. Phys. **10**, 377 (1922).

2. *S. Alexeyev, E. Pamyatnykh, A. Ursulov et al.*, *General theory of relativity: Introduction. Modern development and applications*, URSS publishing group Moscow, 400 (2022).

3. *S. Capozziello, M. De Laurentis*, Phys.Rept. **509**, 167 (2011).

4. *E. Berti, E. Barausse, V. Cardoso, L. Gualtieri, P. Pani*, Class. Quant. Grav. **32**, 243001 (2015).

5. *L. Barack et al*, Class. Quant. Grav. **36**, 143001 (2019)

6. *S. Alexeyev, V. Prokopov*, Universe **8**, 283 (2022)

7. *G. Horndeski*, Int. J. Theor. Phys. **10**, 363 (1974).

8. *T. Kobayashi*, Rept. Prog. Phys. **82**, 086901 (2019).

9. *J. Ezquiaga, M. Zumalac´arregui*, Phys. Rev. Lett. **119**, 251304 (2017).

10. *P. Creminelli, F. Vernizzi*, Phys. Rev. Lett. **119**, 251302 (2017).

11. *A. Starobinsky*, Phys. Lett. B **91**, 99 (1980); Adv. Ser. Astrophys. Cosmol. **3**, 130 (1987).

12. *Y. Ageeva, P. Petrov, V. Rubakov*, Phys. Rev. D **104**, 063530 (2021).

13. *C. Charmousis, E. J. Copeland, A. Padilla, P. M. Saffin*, Phys. Rev. Lett. **108**, 051101 (2012).

14. *E. J. Copeland, A. Padilla, P. M. Saffin*, JCAP **12**, 026 (2012).

15. *Torres, J. C. Fabris, O. F. Piattella*, Phys.Lett.B **798**, 135003 (2019).

16. *X. Calmet, D. Croon, C. Fritz*, Eur. Phys. J. C, **75**, 605 (2015).

17. *S. Alexeyev, X. Calmet, B. Latosh*, Phys. Lett. B, **776**, 111 (2018).

18. *B. Latosh*, Eur. Phys. J. C, **78**, 991 (2018).

19. *B. Latosh*, Eur. Phys. J. C, **80**, 845 (2020).

20. *S. Mironov, V. Rubakov, V. Volkova*, Phys.Rev.D **100**, 083521 (2019).

21. *S. Alexeyev, A. Toporensky, V. Ustiansky*, Class. Quant. Grav **17**, 2243 (2000).

22. *S. Alexeyev, K. Rannu*, JETP **141**, 463 (2012).

23. *S. Alexeyev, M. Senduk*, Universe, **6**, 25 (2020).

24. *P. K. Townsend, P. van Nieuwenhuizen*, Phys.Rev.D **19**, 3592 (1979).

25. *I. Novikov, A. Shatsky, S. Alexeyev, D. Tretyakova*, UFN **184**, 379 (2014).

26. *T. Kobayashi, M. Yamaguchi, J. Yokoyama*, Prog.Theor. Phys. **126** 511 (2011).

27. *S. Sushkov, R. Galeev*, Phys.Rev.D **108** 044028 (2023).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 46–56

NUCLEI, PARTICLES, FIELDS, GRAVITATION, AND ASTROPHYSICS

SPINOR FIELD IN FLRW COSMOLOGY: SPHERICALLY SYMMETRIC CASE

© 2025 B. Saha

Laboratory of Information Technologies, Joint Institute for Nuclear Research, Dubna, Moscow Region, Russia
Peoples' Friendship University of Russia (RUDN University), Moscow, Russia
e-mail: bijan@jinr.ru

Received April 05, 2024

Revised July 31, 2024

Accepted August 30, 2024

Abstract. Within the scope of a spherically symmetric FLRW cosmological model we have studied the role of nonlinear spinor field in evolution of the universe. It is found that if the FLRW model given by the spherical coordinates the energy-momentum tensor (EMT) of the spinor field possesses nontrivial non-diagonal components. These non-diagonal components of EMT neither depend on the spinor field nonlinearity nor on the value of parameter k defining the type of curvature of the FLRW model. The presence of such components imposes some restrictions on the spinor field. The problem is studied for open, flat and close geometries. In doing so we exploited the spinor description of sources such as perfect fluid and dark energies. Some qualitative numerical solutions are given.

DOI: 10.31857/S00444510250104e4

1. INTRODUCTION

Thanks to its ability to simulate different kinds of matter such as perfect fluid, dark energy etc. spinor field is being used by many authors not only to describe the late time acceleration of the expansion, but also to study the evolution of the Universe at different stages [1, 2, 3, 4, 5, 6, 7, 8].

It was found that the spinor field is very sensitive to spacetime geometry. Depending on the concrete type of metric the spinor field may possess different type of nontrivial non-diagonal components of the energy-momentum tensor. As a result the spinor field imposes various kinds of restrictions on both the spacetime geometry and the spinor field itself [9].

Recently spinor field is used in astrophysics to see whether its specific behavior can shed any new light in the study of objects like black hole and wormhole. Such studies were carried out within the scope of spherically symmetric [10, 11] and cylindrically symmetric spacetime [12, 13].

Since the present-day universe is surprisingly isotropic and the presence of nontrivial non-diagonal components of the spinor field leads to the severe restrictions on the spinor field, we have studied role of a spinor field in Friedmann–Lemaitre–Robertson–Walker (FLRW) model as well. But in those cases the space-time was given in Cartesian coordinates. In order to see influence of the coordinate transformations on spinor field some works were done by us earlier [14, 15]. In this paper we will further develop those studies and see how the spinor field behaves if the isotropic and homogeneous cosmological FLRW model given by spherical coordinates.

2. BASIC EQUATION

The action we choose in the form

$$S = \int \sqrt{-g}\left[\frac{R}{2\kappa} + L_{sp}\right] d\Omega, \tag{1}$$

where $\kappa = 8\pi G$ is Einstein's gravitational constant, R is the scalar curvature and L_{sp} is the spinor field Lagrangian given by [16]

$$L_{sp} = \frac{i}{2}[\bar{\psi}\gamma^{\mu}\nabla_{\mu}\psi - \nabla_{\mu}\bar{\psi}\gamma^{\mu}\psi] - m\bar{\psi}\psi - \lambda F(K). \tag{2}$$

To maintain the Lorentz invariance of the spinor field equations the nonlinear term $F(K)$ in (2) is

constructed as some arbitrary functions of invariants generated from the real bilinear forms. On account of Fierz equality in (2) we set $K = K(I,J) = b_1 I + b_2 J$, where b_1 and b_2 takes the value 0 or 1 which leads to the following expressions for $K = \{I,J,I+J,I-J\}$. Here $I = S^2$ and $J = P^2$ are the invariants of bilinear spinor forms with $S = \bar{\psi}\psi$ and $P = i\bar{\psi}\bar{\gamma}^5\psi$ being the scalar and pseudo-scalar, respectively. In (2) λ is the self-coupling constant. Note that λ can be both positive and negative, while $\lambda = 0$ leads to linear case. Here m is the spinor mass.

The covariant derivatives of spinor field takes the form [16]

$$\nabla_\mu \psi = \partial_\mu \psi - \Omega_\mu \psi, \quad \nabla_\mu \bar{\psi} = \partial_\mu \bar{\psi} + \bar{\psi}\Omega_\mu, \qquad (3)$$

where Ω_μ is the spinor affine connections, defined as [16]

$$\Omega_\mu = \frac{1}{4} g_{\rho\sigma}\left(\partial_\mu e^{(b)}_\tau e^\rho_{(b)} - \Gamma^\rho_{\mu\tau}\right)\gamma^\sigma\gamma^\tau. \qquad (4)$$

In (4) $\Gamma^\beta_{\mu\alpha}$ is the Christoffel symbol and the Dirac matrices in curve space–time γ are connected to the flat space–time Dirac matrices $\bar{\gamma}$ in the following way

$$\gamma_\beta = e^{(b)}_\beta \bar{\gamma}_b, \quad \gamma^\alpha = e^\alpha_{(a)}\bar{\gamma}^a, \qquad (5)$$

where $e^\alpha_{(a)}$ and $e^{(b)}_\beta$ are the tetrad vectors such that

$$g_{\mu\nu}(x) = e^a_\mu(x) e^b_\nu(x)\eta_{ab}, \qquad (6)$$

and fulfil following relations

$$e^\alpha_{(a)} e^{(a)}_\beta = \delta^\alpha_\beta, \quad e^\alpha_{(a)} e^{(b)}_\alpha = \delta^b_a. \qquad (7)$$

Here $\eta_{ab} = \mathrm{d}iag(1,-1,-1,-1)$ is the Minkowski spacetime. The γ matrices obey the following anti-commutation rules

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2g_{\mu\nu}, \quad \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}. \qquad (8)$$

Varying the Lagrangian (2) with respect to $\bar{\psi}$ and ψ, respectively, we obtain the following spinor field equations

$$i\gamma^\mu\nabla_\mu\psi - m\psi - D\psi - iG\bar{\gamma}^5\psi = 0, \qquad (9)$$

$$i\nabla_\mu\bar{\psi}\gamma^\mu + m\bar{\psi} + D\bar{\psi} + iG\bar{\psi}\bar{\gamma}^5 = 0, \qquad (10)$$

where $D = 2\lambda F_K b_1 S, \quad G = 2\lambda F_K b_2 P$.

The energy momentum tensor of the spinor field is defined in the following way [16]

$$T_\mu{}^\rho = $$

$$= \frac{i}{4} g^{\rho\nu}(\bar{\psi}\gamma_\mu\nabla_\nu\psi + \bar{\psi}\gamma_\nu\nabla_\mu\psi - \nabla_\mu\bar{\psi}\gamma_\nu\psi - \nabla_\nu\bar{\psi}\gamma_\mu\psi) -$$

$$-\delta^\rho_\mu L, \qquad (11)$$

which in view of (3) we rewrite as

$$T_\mu{}^\rho = $$

$$= \frac{i}{4} g^{\rho\nu}(\bar{\psi}\gamma_\mu\partial_\nu\psi + \bar{\psi}\gamma_\nu\partial_\mu\psi - \partial_\mu\bar{\psi}\gamma_\nu\psi - \partial_\nu\bar{\psi}\gamma_\mu\psi) -$$

$$-\frac{i}{4} g^{\rho\nu}\bar{\psi}(\gamma_\mu\Omega_\nu + \Omega_\nu\gamma_\mu + \gamma_\nu\Omega_\mu + \Omega_\mu\gamma_\nu)\psi - \delta^\rho_\mu L. \qquad (12)$$

Note that the non-diagonal components of the EMT arises thanks to the second term in (12). Moreover, let us emphasize that in view of the spinor field equations (9)–(10) the spinor field Lagrangian (2) can be expressed as

$$L = \lambda\left(2KF_K - F\right), \quad F_K = dF\,/\,dK. \qquad (13)$$

We exploit this form of Lagrangian in solving Einstein equations, as they should be consistent with the Dirac one, as (13) is valid only when spinor fields obey Dirac equations (9)–(10). Let us also note that in case $F = \sqrt{K}$ the Lagrangian vanishes which is very much expected as in this case spinor field becomes linear. We are interested in nonlinear spinor field as only it can generate different kinds of source fields.

The isotropic and homogeneous cosmological model proposed by Friedmann, Lemaitre, Robertson and Walker independently is the most popular and thought to be realistic one among the cosmologists. Let us consider the FLRW model in spherical coordinates in its stanard form [17]:

$$ds^2 = dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\vartheta^2 + r^2\sin^2\vartheta d\phi^2\right], \qquad (14)$$

with k taking the values $+1$, 0 and -1 which corresponds to a close, flat and open universe, respectively. Though the value of k defines the type of geometry of space-time, in reality it is defined by the contents that filled universe. As we see later, independ to the value of k the universe filled with dark energy is always open, whereas for perfect fluid the value of k really matters. In this case depending on the value of k we obtain close, flat or open universe.

In view of (6) the tetrad we will choose in the form

$$e_0^{(0)} = 1, \quad e_1^{(1)} = \frac{a}{\sqrt{1-kr^2}},$$

$$e_2^{(2)} = ar, \quad e_3^{(3)} = ar\sin\vartheta.$$

Then from (5) we find the following γ matrices

$$\gamma^0 = \bar{\gamma}^0, \quad \gamma^1 = \frac{\sqrt{1-kr^2}}{a}\bar{\gamma}^1,$$

$$\gamma^2 = \frac{\bar{\gamma}^2}{ar}, \quad \gamma^3 = \frac{\bar{\gamma}^3}{ar\sin\vartheta}.$$

Further from $\gamma_\mu = g_{\mu\nu}\gamma^\nu$ one finds the γ_μ as well.

The Christoffel symbols, Ricci tensor and scalar curvature and the Einstein tensor corresponding to the metric (14) are well known and can be found in [17].

Then from (4) we find the following expressions for spinor affine connection

$$\Omega_0 = 0, \tag{15}$$

$$\Omega_1 = \frac{1}{2\sqrt{1-kr^2}}\dot{a}\bar{\gamma}^1\bar{\gamma}^0, \tag{16}$$

$$\Omega_2 = \frac{1}{2}r\dot{a}\bar{\gamma}^2\bar{\gamma}^0 + \frac{1}{2}\sqrt{1-kr^2}\bar{\gamma}^2\bar{\gamma}^1, \tag{17}$$

$$\Omega_3 = \frac{1}{2}\dot{a}r\sin\vartheta\bar{\gamma}^3\bar{\gamma}^0 + \frac{1}{2}\sqrt{1-kr^2}\sin\vartheta\bar{\gamma}^3\bar{\gamma}^1 + \frac{1}{2}\cos\vartheta\bar{\gamma}^3\bar{\gamma}^2. \tag{18}$$

Let us consider the case when the spinor field depends on t only, then in view of (15)–(18) the spinor field equations can be written as

$$\dot{\psi} + \frac{3}{2}\frac{\dot{a}}{a}\psi + \frac{\sqrt{1-kr^2}}{ar}\bar{\gamma}^0\bar{\gamma}^1\psi + \frac{\cot\vartheta}{2ar}\bar{\gamma}^0\bar{\gamma}^2\psi + i(m+D)\bar{\gamma}^0\psi + G\bar{\gamma}^5\bar{\gamma}^0\psi = 0, \tag{19}$$

$$\dot{\bar{\psi}} + \frac{3}{2}\frac{\dot{a}}{a}\bar{\psi} - \frac{\sqrt{1-kr^2}}{ar}\bar{\psi}\bar{\gamma}^0\bar{\gamma}^1 - \frac{\cot\vartheta}{2ar}\bar{\psi}\bar{\gamma}^0\bar{\gamma}^2 - i(m+D)\bar{\psi}\bar{\gamma}^0 + G\bar{\psi}\bar{\gamma}^5\bar{\gamma}^0 = 0, \tag{20}$$

Introducing $\varphi = a^{3/2}\psi$ we rewrite the equation (19)–(20)

$$\dot{\varphi} + \frac{\sqrt{1-kr^2}}{ar}\bar{\gamma}^0\bar{\gamma}^1\varphi + \frac{\cot\vartheta}{2ar}\bar{\gamma}^0\bar{\gamma}^2\varphi + i(m+D)\bar{\gamma}^0\varphi + G\bar{\gamma}^5\bar{\gamma}^0\varphi = 0, \tag{21}$$

$$\dot{\bar{\varphi}} - \frac{\sqrt{1-kr^2}}{ar}\bar{\varphi}\bar{\gamma}^0\bar{\gamma}^1 - \frac{\cot\vartheta}{2ar}\bar{\varphi}\bar{\gamma}^0\bar{\gamma}^2 - i(m+D)\bar{\varphi}\bar{\gamma}^0 + G\bar{\varphi}\bar{\gamma}^5\bar{\gamma}^0 = 0, \tag{22}$$

The equation (21) can be presented in the matrix form

$$\dot{\varphi} = A\varphi, \tag{23}$$

or

$$\begin{pmatrix}\dot{\varphi}_1\\ \dot{\varphi}_2\\ \dot{\varphi}_3\\ \dot{\varphi}_4\end{pmatrix} = \begin{pmatrix}-iD_1 & 0 & -G & B_1\\ 0 & -iD_1 & B_1^* & -G\\ G & B_1 & iD_1 & 0\\ B_1^* & G & 0 & iD_1\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\\ \varphi_3\\ \varphi_4\end{pmatrix}, \tag{24}$$

where

$$D_1 = (m+D), \quad B_1 = -\frac{\sqrt{1-kr^2}}{ar} + i\frac{\cot\vartheta}{2ar},$$

$$B_1^* = -\frac{\sqrt{1-kr^2}}{ar} - i\frac{\cot\vartheta}{2ar}.$$

It can be shown that

$$detA = \left(D_1^2 + G^2 - B_1B_1^*\right)^2.$$

We can choose the nonlinearity in such a way that the corresponding determinant is nontrivial. In that case the solution (23) can be formally written as [18]

$$\varphi(t) = T\exp\left(-\int_t^{t_1} A_1(\tau)d\tau\right), \tag{25}$$

where $T = \varphi(t_1)$ is the solution at $t = t_1$. Given the fact that the universe is expanding and the spinor field invariants are the inverse functions of scale factor, in case of a nonzero spinor mass one can assume

$$\varphi(t_1) = col\left(\varphi_1^0 e^{-imt_1}, \varphi_2^0 e^{-imt_1}, \varphi_3^0 e^{imt_1}, \varphi_4^0 e^{imt_1}\right),$$

whereas for a massless spinor field

$$\varphi(t_1) = col\left(\varphi_1^0, \varphi_2^0, \varphi_3^0, \varphi_4^0\right)$$

with φ_i^0 being constants.

The non-trivial components of the energy momentum tensor of the spinor field in this case read

$$T_0^0 = mS + \lambda F, \tag{26}$$

$$T_1^1 = T_2^2 = T_3^3 = -\lambda\left(2KF_K - F\right), \tag{27}$$

$$T_3^1 = \frac{a\cos\vartheta}{4\sqrt{1-kr^2}}A^0, \tag{28}$$

$$T_1^0 = \frac{\cot\vartheta}{4r\sqrt{1-kr^2}}A^3, \tag{29}$$

$$T_2^0 = -\frac{3}{4}\sqrt{1-kr^2}\,A^3, \tag{30}$$

$$T_3^0 = \frac{3}{4}\sqrt{1-kr^2}\sin\vartheta A^2 - \frac{1}{2}\cos\vartheta A^1. \tag{31}$$

From (28)–(31) we conclude that the energy-momentum tensor of the spinor field contains nontrivial non-diagonal components. The non-diagonal components

- do not depend on the spinor field nonlinearity;
- occur due to the spinor affine connections;
- appear depending on space-time geometry as well as the system of coordinates;
- impose restrictions on spinor field and/or space-time geometry;
- do not depend on the value of k which defines the type of curvature.

It should be emphasized that for a FLRW model given in Cartesian coordinate the EMT have only diagonal components with all the non-diagonal one being identically zero [19]. So in this case the non-diagonal components arise as a result of coordinate transformation. Note also that all cosmological spacetime defined by diagonal matrices of Bianchi type VI, VI_0, V, III, I, $LRS-BI$ and $FLRW$, possess same diagonal components of EMT, but has nontrivial non-diagonal elements that differ from each other in different cases [9]. Moreover, non-diagonal metrics such as Bianchi type II, $VIII$ and IX also have nontrivial non-diagonal components of EMT. Consequently, we see that the appearance of non-diagonal components of the energy-momentum tensor occurs either due to coordinate transformations or due to the geometry of space-time.

As one sees, the components of the EMT of the spinor field contains some spinor field invariants. To define those invariants let us write the system of equations for the invariants of the spinor field. It can be obtained from the spinor field equation (19)–(20):

$$\dot{S}_0 + 2GA_0^0 = 0, \tag{32}$$

$$\dot{P}_0 - 2\left(m + D\right)A_0^0 = 0, \tag{33}$$

$$\dot{A}_0^0 + 2GS_0 + 2\left(m + D\right)P_0 + 2\frac{\sqrt{1-kr^2}}{ar}A_0^1 + \frac{\cot\vartheta}{ar}A_0^2 = 0, \tag{34}$$

$$\dot{A}_0^1 + 2\frac{\sqrt{1-kr^2}}{ar}A_0^0 = 0, \tag{35}$$

$$\dot{A}_0^2 + \frac{\cot\vartheta}{ar}A_0^0 = 0, \tag{36}$$

that gives the following relation between the invariants:

$$P_0^2 - S_0^2 + \left(A_0^0\right)^2 - \left(A_0^1\right)^2 - \left(A_0^2\right)^2 = C_0, C_0 = \text{const}. \tag{37}$$

In (32)–(37) the quantities with a subscript "0" are related to the normal ones as follows: $X_0 = Xa^3$. From (37) we can conclude that since C_0 is an arbitrary constant, the each term of (37) should be constant as well.

In order to solve the Einstein equations we have to know how the components of the EMT are related to the metric functions. In order to know that let us find the invariant K in general. We consider the 4 cases separately.

In case of $K = I$, $G = 0$. In this case from (32) we find

$$S = \frac{C_s}{a^3}, \quad \Rightarrow K = \frac{C_s^2}{a^6}. \tag{38}$$

If $K = J$, then in case of a massless spinor field from (33) we find

$$P = \frac{C_p}{a^3}, \quad \Rightarrow K = \frac{C_p^2}{a^6}. \tag{39}$$

Let us consider the case when $K = I + J$. In this case $b_1 = b_2 = 1$. Then on account of expression for D and G from (32) and (33) for the massless spinor field we find

$$\dot{S}_0 + 4\lambda a^3 F_K P A^0 = 0, \tag{40}$$

$$\dot{P}_0 - 4\lambda a^3 F_K S A^0 = 0, \tag{41}$$

which yields

$$K = I + J = S^2 + P^2 = \frac{C_1^2}{a^6}. \tag{42}$$

Finally in case when $K = I - J$, i.e. $b_1 = -b_2 = 1$ from (32) and (33) for the massless spinor field we find

$$\dot{S}_0 + 4\lambda a^3 F_K P A^0 = 0, \tag{43}$$

$$\dot{P}_0 + 4\lambda a^3 F_K S A^0 = 0, \quad (44)$$

which yields

$$K = I - J = S^2 - P^2 = \frac{C_2^2}{a^6}. \quad (45)$$

Thus we see that the invariant K is a function of metric function a, namely, $K = const.a^{-6}$ and it is what we need to solve the Einstein equation. In what follows we solve the Einstein equation.

Let us recall that the Einstein tensor G_μ^ν corresponding to the metric (14) possesses only nontrivial diagonal components. Hence the general Einstein system of equations

$$G_\mu^\nu = -8\pi G\, T_\mu^\nu, \quad (46)$$

leads to the following non-diagonal expressions

$$0 = T_\mu^\nu, \quad \mu \neq \nu. \quad (47)$$

In view of (28)–(31) from (47) one dully finds that

$$A^0 = 0, \quad A^3 = 0, \quad A^1 = (3/2)\sqrt{1 - kr^2}\, \vartheta A^2. \quad (48)$$

Note that since the FLRW model given by the Cartesian coordinate the non-diagonal components of EMT are identically zero, hence relation such as (48) does not exist.

In view of $A^0 = 0$, $A^3 = 0$ from the system (32)–(36) we find

$$S_0 = C_S, \quad P_0 = C_P, \quad A_0^1 = C_0^1, \quad A_0^2 = C_0^2, \quad (49)$$

with C_S, C_P, C_0^1 and C_0^2 being some arbitrary constants. Thus we see that $K = const.a^{-6}$. Note that the equation (34) in this case in redundant and (48) gives relations between the constants C_0^1 and C_0^2.

We are now ready to consider the diagonal components of the Einstein system of equations which for the metric (14) takes the form

$$2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) = 8\pi G T_1^1, \quad (50)$$

$$3\left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) = 8\pi G T_0^0. \quad (51)$$

On account of (51) we rewrite (50) in the form

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(T_0^0 - 3T_1^1\right) = -\frac{4\pi G}{3}(\varepsilon + 3p), \quad (52)$$

where ε and p are the the energy density and and pressure, respectively:

$$\varepsilon = T_0^0 = mS + \lambda F, \quad (53)$$

$$p = -T_1^1 = \lambda(2KF_K - F). \quad (54)$$

On account of (26) and (27) from (52) we find

$$\ddot{a} = -\frac{4\pi G}{3}(mS - 2\lambda F + 6\lambda K F_K)a. \quad (55)$$

Note that the equations (52) or (55) do not contain k that defines the type of space-time curvature. In order to take this very important quantity into account we have to exploit (51) as the initial condition for $\dot{a}$. The equation (51) we rewrite in the form

$$\dot{a} = \pm\sqrt{(8\pi/3)G\,\varepsilon a^2 - k} = \\ = \pm\sqrt{(8\pi/3)G\,(mS + \lambda F)a^2 - k}, \quad (56)$$

Now we can solve (55) with the initial condition given by (56). It comes out that these equations are consistent when one takes the negative sign in (56). Alternatively, one can solve (56), but for the system to be consistent he has to check whether the result satisfies (55).

As we have already established, S, K, hence $F(K)$ are the functions of a. Consequently, given the spinor field nonlinearity the foregoing equation can be solved either analytically or numerically.

The equation (55) can be solved analytically. The first integral of (55) takes the form

$$\dot{a} = \sqrt{\int f(a)da + C_c}, \quad (57)$$

where we define

$$f(a) = -\frac{8\pi G}{3}(mS - 2\lambda F + 6\lambda K F_K)a$$

and C_c is a constant which should be defined from (56). The solution to the equation (57) can be given in quadrature

$$\int \frac{da}{\sqrt{\int f(a)da + C_c}} = t. \quad (58)$$

1. In what follows we solve the system (50)–(51) numerically. In doing so we rewrite it in the following way:

$$\dot{a} = Ha, \quad (59)$$

$$\dot{H} = -\frac{3}{2}H^2 - \frac{1}{2}\frac{k}{a^2} - 4\pi G\lambda\left(2KF_K - F\right), \quad (60)$$

$$H^2 = \frac{8\pi G}{3}\left(mS + \lambda F\right) - \frac{k}{a^2}, \quad (61)$$

where H is the Hubble constant.

As one sees, in the foregoing system the first two are differential equations, whereas the third one is a constraint, which we use as the initial condition for H :

$$H = \pm\sqrt{8\pi G\left(mS + \lambda F\right)/3 - k/a^2}. \quad (62)$$

Since the expression under the square-root must be non-negative, it imposes some restrictions on the choice of initial value of a as well. Note that initial value of H depends on spinor mass m , coupling parameter λ and the value of k .

3. NUMERICAL SOLUTIONS

In what follows we solve the equations (59) and (60), numerically. The third equation of the system (61) we exploit as initial condition for $H(t)$ in the form (62). We do it for both massive and massless spinor field. Beside this, we consider close, flat and open universe choosing different values for k . As it was mentioned earlier, the coupling constant λ can be positive or negative. Let us recall that

$$K = \frac{K_0}{a^6}, \quad K_0 = \text{const}. \quad (63)$$

The foregoing relation holds for $K = \{I, J, I \pm J\}$ for a massless spinor field, whereas for $K = I = S^2$ it is true for both massive and massless spinor field. Hence we assume that $K = I = S^2$. We consider different kind of spinor field nonlinearities $F(K)$ (equivalently, $F(S)$), that describes various types of sources from perfect fluid to dark energy.

3.1 Barotropic equation of state

Let us consider the case when the Universe is filled with perfect fluid or dark energy given by quintessence, Λ -term or phantom matter. It can be implemented by the barotropic equation of state (EoS), which gives a linear dependence between the pressure and energy density and was exploited by many authors [20, 21, 22, 23]. The corresponding EoS takes the form

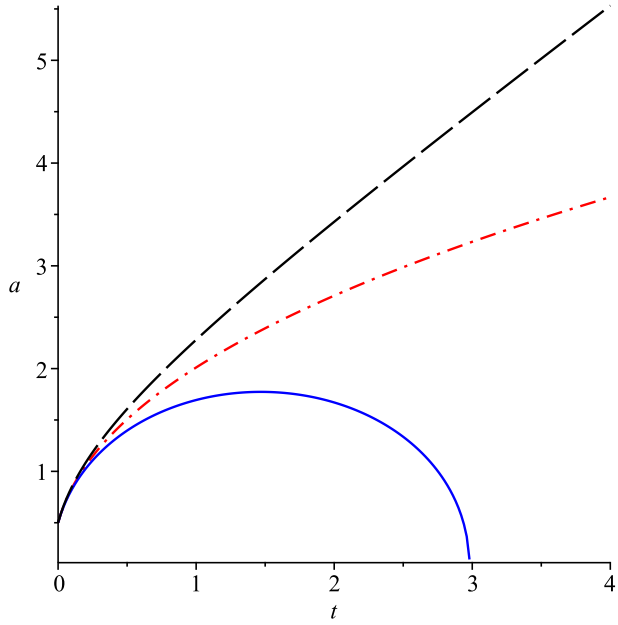

Fig. 1. Evolution of the FLRW Universe (scale factor $a(t)$) in presence of a radiation given by a massless spinor field. The blue solid, red dash-dot and black long dash lines stand for close, flat and open ($k = +1, 0, -1$) universe, respectively

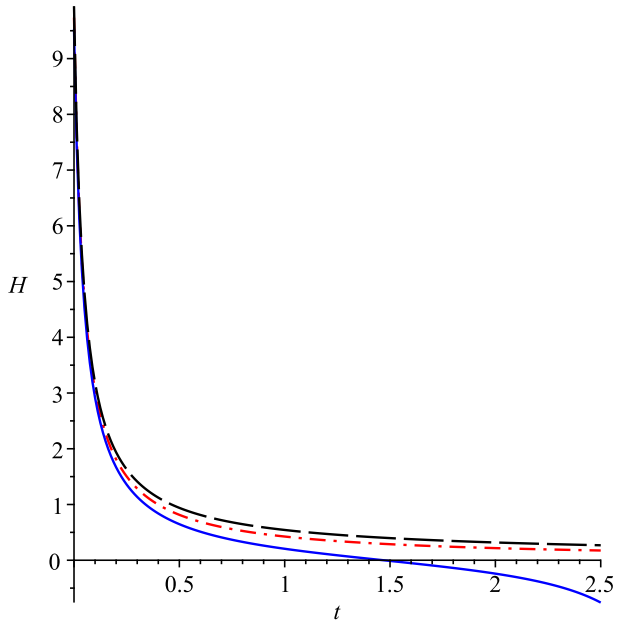

Fig. 2. Evolution of the corresponding Hubble parameter $H(t)$ and corrsponds to differnt values of k as in Fig. 1

$$p = W\varepsilon, \quad (64)$$

where the EoS parameter W is a constant. Depending on the value of W, the Eq. (64) can give rise to both perfect fluid, such as dust, radiation etc. and dark

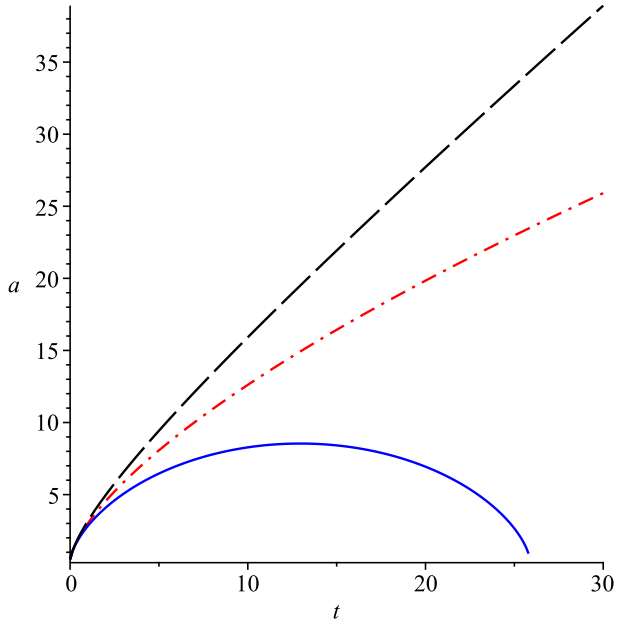

Fig. 3. Evolution of the FRW Universe (scale factor $a(t)$) in presence of a radiation given by a massive spinor field. The blue solid, red dash-dot and black long dash lines stand for $k = +1, 0, -1$, respectively

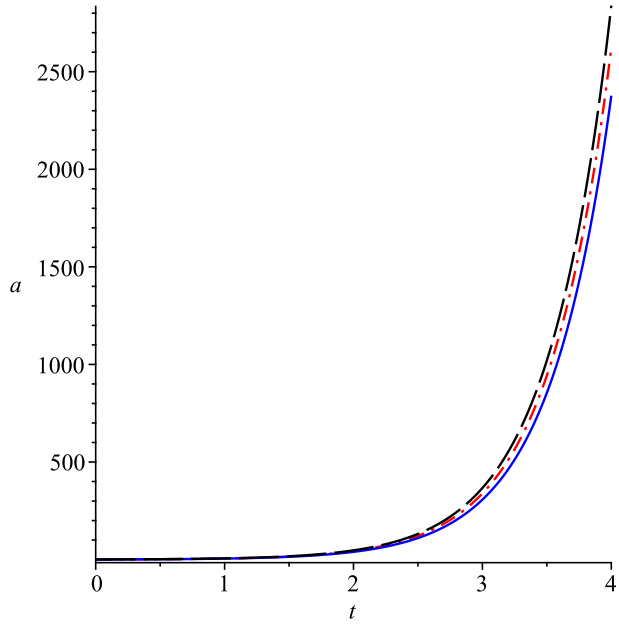

Fig. 5. Evolution of the FRW Universe (scale factor $a(t)$) in presence of a modified Chaplygin gas given by a massless spinor field. As one sees, independent to the value of k in this case the universe expand rapidly

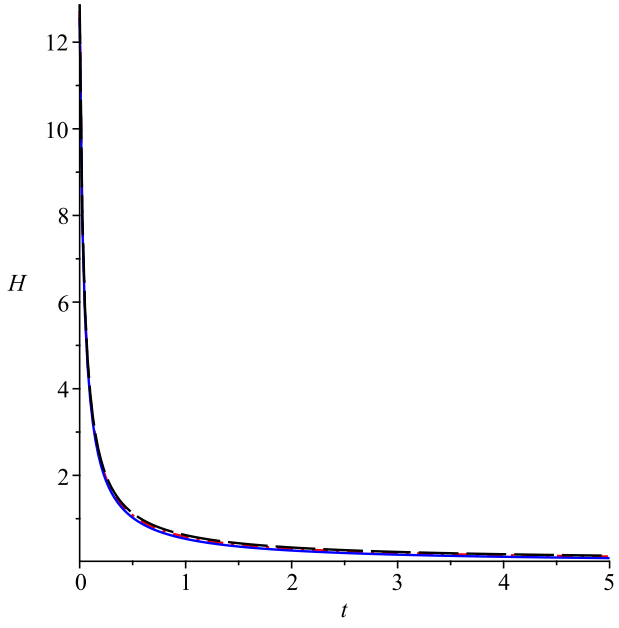

Fig. 4. Evolution of the corresponding Hubble parameter $H(t)$

energy such as quintessence, cosmological term, phantom matter etc. For $W \in [0,1]$, it describes a perfect fluid. The value $W = -1$ represents a typical cosmological constant (Λ-term) [24, 25, 26], whereas $W \in [-1,-1/3]$ gives rise to a quintessence, while for $W < -1$ it ascribes a phantom matter.

It was shown in [9, 27] that inserting (26)–(27) into (64) the matter or energy corresponding to Eq. (64) can be simulated by the nonlinear term given by

$$F(S) = \lambda S^{1+W} - mS, \quad \lambda = \text{const.}, \qquad (65)$$

in the spinor field Lagrangian (2).

Let us now solve (59)–(61) numerically for the nonlinear term given by (65). We consider both massive and massless spinor field. The values of W are taken to be $1/2$, $-1/2$ and -1 describing the radiation, quintessence and cosmological constant, respectively. For simplicity we set $S_0 = 1$, $G = 1$, $\lambda = 0.5$ here and in the cases to follow. We also set $m = 0$ for a massless and $m = 1$ for a massive spinor field.

In Fig. 1 we have illustrated the evolution of the Universe filled with radiation, given by a massless spinor field, while Fig. 2 shows the evolution of the Hubble parameter corresponding to the case in question. Figs. 3 and 4 describes the evolution of the Universe filled with radiation and the corresponding Hubble parameter in case of a massive spinor field. In the figures blue solid line stands for a closed universe given by $k = 1$, red dash-dot line stands for a flat universe with $k = 0$ and black long dash line stands for an open universe with $k = -1$.

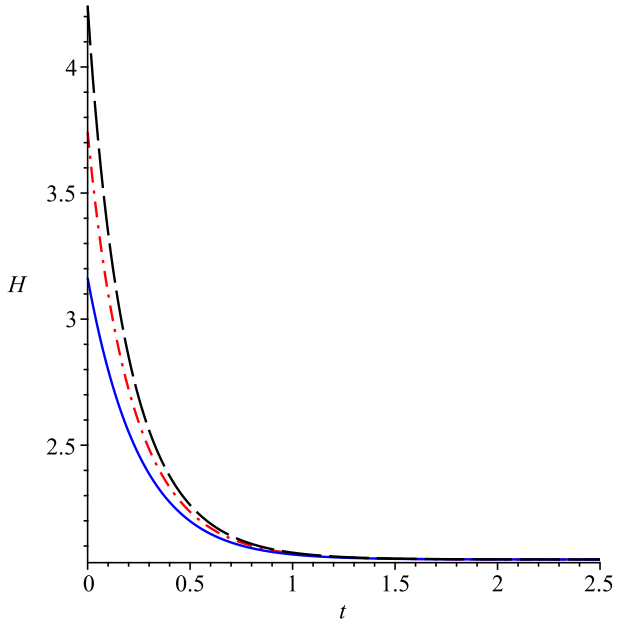

Fig. 6. Evolution of the corresponding Hubble parameter $H(t)$

We have also considered the case with the spinor field nonlinearity describing a quintessence $(W = -1/2)$ and cosmological constant $(W = -1)$. Both massive and massless spinor fields are taken into account. Since in both cases the energy density is less than the critical density, independent to the value of k we have only open type of universe. The behavior of the evolution is qualitatively same as that of in case of a modified Chapligin gas. The corresponding figures will be similar to those in Figs. 5 and 6, only the rate of expansion being much slower.

3.2 Chaplygin gas

In order to combine two different physical concepts such as dark matter and dark energy, and thus reduce the two physical parameters in one, a rather exotic equation of state was proposed in [28] which was further generalized in the works [29, 30]. Generalized Chaplygin gas model is given by the EoS

$$p_{ch} = -A/\varepsilon_{ch}^{\alpha}, \qquad (66)$$

where A is a positive constant and $0 < \alpha \leq 1$.

It was shown that such kind of dark energy can be modeled by the massless spinor field with the nonlinearity [9] inserting (26)–(27) into (66)

$$F(S) = \left(A + \lambda S^{1+\alpha}\right)^{1/(1+\alpha)}. \qquad (67)$$

We have solved (59)–(61) numerically for the nonlinear term given by (67). We consider only massless spinor field setting $m = 0$. The parameters S_0, G and λ were taken as in previous case. We have also set $A = 1/2$ and $\alpha = 1/3$.

As in case of quintessence and cosmological constant, the evolution of the universe filled with Chaplygin gas and corresponding behavior of the Hubble parameter are qualitatively same as in case of a modified Chaplygin gas which are illustrated in Figs. 5 and 6. The expansion rate in this case is higher than the previous case but slower than in the case to follow.

3.3 Modified Chaplygin gas

Though the dark energy and the dark matter act in a completely different way, many researchers suppose that they are different manifestations of a single entity. Following such an idea a modified Chaplygin gas was introduced in [31] and was further developed in [32]. Corresponding EoS takes the form

$$p = W\varepsilon - A/\varepsilon^{\alpha}, \qquad (68)$$

with W being a constant, $A > 0$ and $0 \leq \alpha \leq 1$.

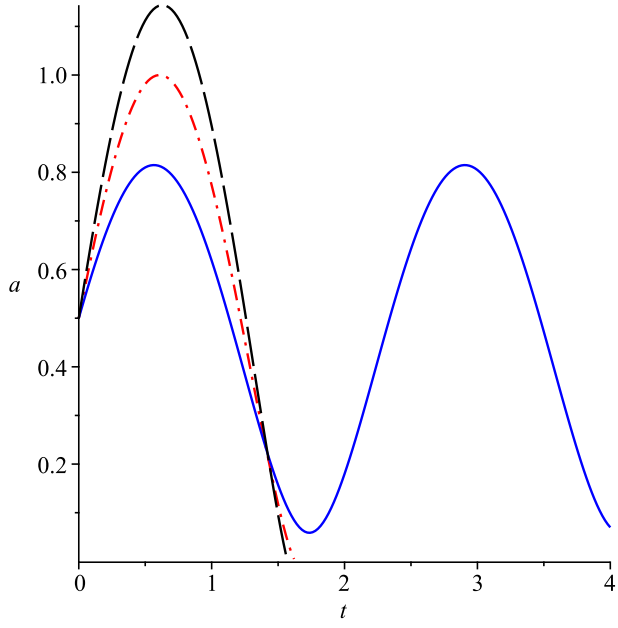

Fig. 7. Evolution of the FRW Universe (scale factor $a(t)$ in presence of a modified quintessence given by a massless spinor field. In case of $k = +1$ there occurs a periodic solution, whereas for $k = 0$ or $k = -1$, we have Big Crunch like solutions

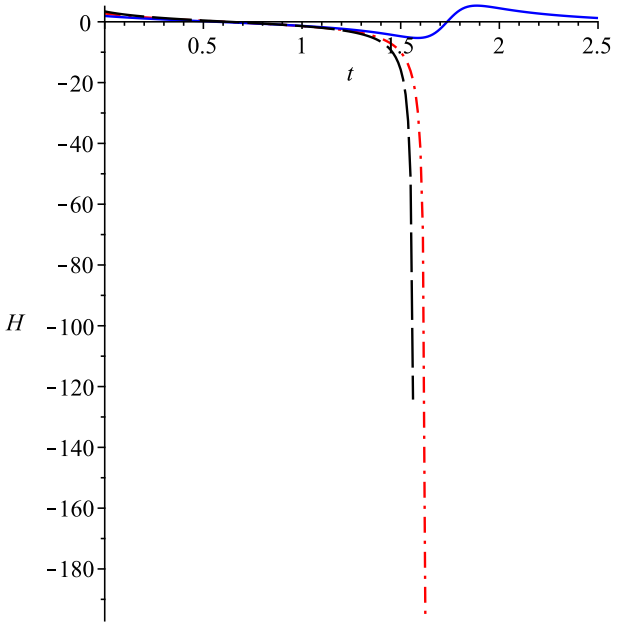

Fig. 8. Evolution of the corresponding Hubble parameter $H(t)$

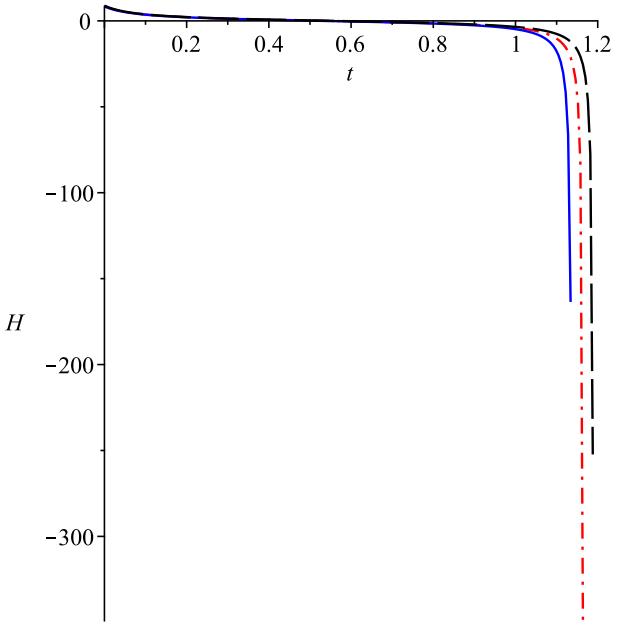

Fig. 10. Evolution of the corresponding Hubble parameter $H(t)$

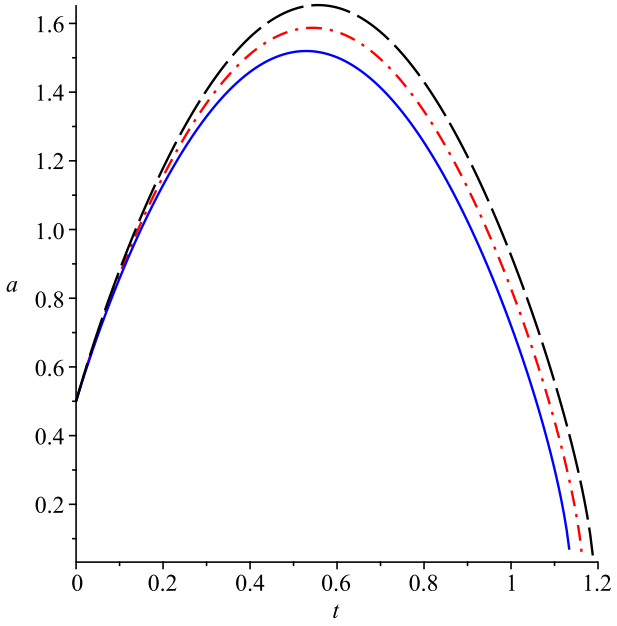

Fig. 9. Evolution of the FRW Universe (scale factor $q(t)$) in presence of a modified quintessence given by a massive spinor field. Unlike massless spinor field, in this case there is no periodic solutions for the given value of problem parameters

Inserting (26)–(27) into (68) the modified Chaplygin gas can be generated by a massless spinor field with the nonlinearity given by [9]

$$F(S) = \left[\frac{A}{1+W} + \lambda S^{(1+\alpha)(1+W)}\right]^{1/(1+\alpha)}. \qquad (69)$$

In fact, mathematically it is a combination of quintessence and Chaplygin gas. We have solved (59)–(61) numerically for the nonlinear term given by (69). Since we consider only massless spinor field, we set $m = 0$. For simplicity we set S_0, G, λ, A, and α as in previous cases. Beside that we set $W = -1/2$.

In Figs. 5 and **6** we have illustrated the evolution of the universe and corresponding Hubble parameter when the Universe is filled with nonlinear spinor field simulating a modified Chaplygin gas.

3.4 Modified quintessence

A modified Quintessence was proposed in order to avoid eternal acceleration of the universe. In some cases it gives cyclic universe that pops up from a Big Bang singularity, expands to some maximum value and then decreases and finally ends in Big Crunch. In some cases it might be periodic without singularity. A spinor description of a modified quintessence was proposed in [23]

$$p = W(\varepsilon - \varepsilon_{cr}), \quad W \in (-1,0), \qquad (70)$$

with ε_{cr} being some critical energy density. The model gives rise to cyclic or oscillatory universe. Setting $\varepsilon_{cr} = 0$ one obtains ordinary quintessence. As one sees from (70), the pressure is negative as long as $\varepsilon > \varepsilon_{cr}$. Since with the expansion of the universe the energy density decreases, at some moment of

time ε becomes less than ε_{cr}, i.e., $\varepsilon < \varepsilon_{cr}$. This leads to the positive pressure and the contraction of the universe. It can be shown that a modified quintessence can be modeled by a spinor field nonlinearity inserting (26)–(27) into (70)

$$F(S) = \lambda S^{1+W} + \frac{W}{1+W}\varepsilon_{cr}. \qquad (71)$$

We solve the system (59)–(61) for the values of parameters as in case of quintessence. For critical density we set $\varepsilon_{cr} = 1$.

In Figs. 7 and 8 we have illustrated the evolution of the universe and corresponding Hubble parameter when the universe is filled with nonlinear massless spinor field simulating a modified quintessence. The corresponding cases with massive spinor field are illustrated in Figs. 9 and 10

In the figures, evolution of Hubble parameter H is drawn for a much smaller time interval than the scale factor a. It is just for technical reason. For example, if in Figs. 3 and 4 we use interval 30 for both a and H, as we see from Fig. 4 Hubble parameter after crossing mark 5 it becomes almost zero, thus giving rise to a visually ugly picture. Whereas, setting interval 5 for both, we have a on rising phase for all three values of k [cf. Fig. 3]. These two figures correspond to the same values of problem parameter, only for good visual pictures we have drawn them for different intervals. The same can be told for all other cases.

4. CONCLUSION

Within the scope of a spherically symmetric FLRW model we have studied the role of a nonlinear spinor field in the evolution of the universe. It is found that in this case the spinor field possesses nontrivial non-diagonal components of the EMT. Since the Einstein tensor in this case is diagonal, this fact imposes some restrictions on the components of spinor field: $A^0 = 0$, $A^3 = 0$ and $A^1 \propto A^2$. Corresponding equations are solved. It is shown that if the spinor field nonlinearity repesents ordinay matter such as radiation, the factor k plays decisive role giving rise to close, flat or open universe depending on its positive, trivial or negative values. It is also shown that in this case spinor mass influences the result quantatively. If the spinor feild nonlinearity generates a dark energy we have only rapidly expanding universe independent to the value of k. Finally in case of a modified quintessence the model gives rise to as oscillating universe. Depending on the value of k and spinor mass m there might be periodic solutions or the one that ends in Big Crunch.

FUNDING

This paper has been supported by the RUDN University Strategic Academic Leadership Program.

REFERENCES

1. *M. O. Ribas, F. P. Devecchi, and G. M. Kremer*, Phys. Rev. D **72**, 123502 (2005).
2. *B. Saha*, Phys. Rev. D **74**, 124030 (2006).
3. *B. Saha*, Gravitation and Cosmology **12**, 215 (2006).
4. *B. Saha and V. S. Rikhvitsky*, Phys. Part. Nucl. **40**, 612 (2009).
5. *L. Fabbri*, Int. J. Theor. Phys. **52**, 634 (2013).
6. *L. Fabbri*, Phys. Rev. D **85**, 047502 (2012).
7. *S. Vignolo, L. Fabbri, and R. Cianci*, J. Math. Phys. **52**, 112502 (2011).
8. *N. J. Popławski*, Phys. Rev. D8**5**, 107502 (2012).
9. *B. Saha*, Phys. Parti. Nucl. **49**, 146 (2018).
10. *B. Saha,* Eur. Phys. J. Plus **133** 461 (2018).
11. *K. A. Bronnikov, Yu. P. Rybakov, and Saha B. Saha*, Eur. Phys. J. Plus **135** 124 (2020).
12. *B. Saha*, Universe **6**, 152 (2020).
13. *B. Saha*, Eur. Phys. J. Plus **137** 1063 (2022).
14. *B. Saha, E. I. Zakharov, and V. S. Rikhvitsky*, Discrete and Continous Models and Applied Computational Science **28**, 132 (2020).
15. *A. S. Gavrikov, B. Saha, and V. S. Rikhvitsky*, Discrete and Continous Models and Applied Computational Science **28**, 120 (2020).
16. *B. Saha*, Phys. Rev. D **64**, 123501 (2001).
17. *J. V. Narlikar Introduction to Relativity*, Cambridge University Press, NY (2010).
18. *B. Saha*, Eur. Phys. J. Plus **131** 170 (2016).
19. *B. Saha*, AstroPhys. Space Sci. **365**, 68 (2020).
20. *R. R. Caldwell*, Phys. Lett. B **545**, 23 (2002).
21. *V. Sahni and A. A. Starobinsky*, Int. J. Mod. Phys. D **9**, 373 (2000).
22. *I. Zlatev, L. Wang, and P. J. Steinhardt*, Phys. Rev. Lett. **82**, 896 (1999).
23. *B. Saha*, Int. J. Theor. Phys. **45**, 983 (2006).
24. *T. Padmanabhan*, Phys. Reports **380**, 235 (2003).
25. *V. Sahni*, Lecture Notes on Phys. **653**, 141 (2004).
26. *B. Saha*, Astrophys. Space Sci. **302**, 83 (2006).
27. *B. Saha*, Astrophys. Space Sci. **331**, 243 (2011).

28. *A. Yu. Kamenshchik, U. Moschella, and V. Pasquier*, Phys. Lett. B **511**, 265 (2001).

29. *N. Bilic, G. B. Tupper, and R. D. Viollier*, Phys. Lett. B **353**, 17 (2002).

30. *M. C. Bento, O. Bertolami, and A. A. Sen*, Phys. Rev. D **66**, 043507 (2002).

31. *H. B. Benaoum*, Universe **8**, 340 (2022).

32. *H. B. Benaoum*, Adv. High Energy Phys. 2012:357802 (2012).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 57–63

NUCLEI, PARTICLES, FIELDS, GRAVITATION, AND ASTROPHYSICS

THE QUANTUM MODELS OF AN ELECTRON WITH THE ZERO SELF-ENERGY

© 2025 V. P. Neznamov*, V. E. Shemarulin

Russian Federal Nuclear Center – All-Russian Research Institute of Experimental Physics (RFNC–VNIIEF), 607188 Sarov, Nizhny Novgorod Region, Russia

**e-mail: vpneznamov@mail.ru, vpneznamov@vniief.ru*

Received April 21, 2024

Revised September 02, 2024

Accepted September 02, 2024

Abstract. We propose two models of an electron with the zero self-energy based on the updated regular charged Reissner-Nordström and Kerr-Newman metrics with quantum nuclei.

DOI: 10.31857/S00444510250105e6

1. INTRODUCTION

Since the advent of General Relativity (GR), attempts have been made to construct models of elementary particles in curved spacetime. Notable contributors to such models include G.B. Jeffery (1921), P. A.M. Dirac (1962), W. Israel (1970), C.A. López (1984), O. Gron (1984), A. Burinskii (1974–2023), and others. Unfortunately, none of the proposed models have found practical application in classical and quantum field theory calculations.

Another longstanding problem, which has engaged many researchers and is the focus of this paper, is the issue of the infinite self-energy of a charged particle in classical and quantum electrodynamics. Efforts to eliminate the linear divergence of self-energy in classical electrodynamics were made by H. Poincaré, M. Born, L. Infeld, P. A.M. Dirac, J. Wheeler, R. Feynman, and others. In quantum field theory, the renormalization procedure for fermion masses was developed to address the logarithmic divergence of self-energy.

Such efforts continue today. For example, in [1, 2], quantum electrodynamics demonstrates that the self-energy of a point charge converges when the nonlinearity of the theory is considered in any finite order of the Euler–Heisenberg Lagrangian expansion in powers of the electric field.

In this paper, using the electron as an example, we propose two quantum models of charged elementary particles with zero self-energy. By employing the quantum geometry of the Reissner–Nordström (RN) metric and neglecting extremely small gravitational coefficients, all practical calculations in classical and quantum electrodynamics can be conducted within the paradigm of elementary particles as point masses with electric charges.

Our approach is based on the phenomenological description of quantum black holes for modified Schwarzschild (Sq) and Reissner–Nordström (RNq) geometries [3, 4]. In this framework, black holes contain quantum cores described by coherent states of gravitons. The coherent-state-averaged solutions of the massless Klein–Gordon equation for longitudinal gravitons are equated, with certain coefficients, to classical potentials. Short wavelengths are eliminated by a graviton energy cut-off, introducing a maximum graviton energy:

$$k_{UV} = \frac{\hbar c}{R_S}. \tag{1}$$

For convenience, as in [3, 4], we introduce the parameter R_S. The primary quantity in this theory is the maximum graviton energy k_{UV}. The presence of a quantum core gives rise to quantum "hairs." Quantum black holes thus possess quantum hairs.

In a future quantum theory of gravity, the graviton energy cut-off k_{UV} will be replaced by strict integration, and the absence of short wavelengths in graviton coherent states will naturally result from the application of a more advanced quantum theory.

In our previous work [5], we extended the approach of [3, 4] to modified M and Kerr–Newman (KNq) geometries, describing regular uncharged and charged quantum rotating collapsars. As with the RNq geometry, this term includes either black holes with quantum cores and event horizons or rotating quantum cores without event horizons.

In [5], for charged rotating collapsars with mass M, charge Q, and angular momentum J, we obtained full regularization of the KNq quantum metrics at the following parameter value:

$$R_S = R_S^{reg} = \frac{\pi}{8}\frac{Q^2}{Mc^2} \tag{2}$$

This regularization yielded finite values for key GR quantities, such as the mass function $m_{KNq}(r)$, $R_q^{\mu\nu}(r,\theta)$, the Kretschmann scalar $K_q(r,\theta)$, and others.

For $R_S = R_S^{reg}$, the total energy of the quantum charged rotating collapsar equals $E = Mc^2$, meaning its self-energy is zero. Due to the presence of a quantum core, the electromagnetic forces responsible for the collapsar's self-energy are counterbalanced by gravitational forces.

Similar results are obtained for the RNq quantum metric [4].

In Section 2, we propose two quantum electron models with zero self-energy based on RNq and KNq quantum geometries. Section 3 compares these models, favoring the RNq-based electron model. The conclusion summarizes the key findings of this paper.

The Appendix provides the procedure for calculating the energy of a charged rotating black hole with a quantum core (see [5]).

2. QUANTUM MODELS OF THE ELECTRON

Based on regular quantum models of charged rotating and non-rotating black holes [4, 5], we propose two quantum models of the electron with modified KNq and RNq metrics.

2.1 Modified Kerr–Newman geometry

For the electron model, we will use the Cürses-Cürsey metric [6][1]:

$$ds^2_{KNq} = \left(1 - \frac{2r m^e_{KNq}(r)}{\rho^2}\right)dt^2 + \frac{4a_e r m^e_{KNq}(r)\sin^2\theta}{\rho^2}dtd\varphi - \frac{\rho^2}{\Delta}dr^2 - \rho^2 d\theta^2 - \frac{\Sigma\sin^2\theta}{\rho^2}d\varphi^2, \tag{3}$$

where $m^e_{KNq}(r)$ is the mass function,

$$\rho^2 = r^2 + a_e^2\cos^2\theta, \tag{4}$$

$$\Delta = r^2 - 2r m^e_{KNq}(r) + a_e^2, \tag{5}$$

$$\Sigma = \left(r^2 + a_e^2\right)^2 - a_e^2\Delta\sin^2\theta, \tag{6}$$

$$a_e = \frac{|J_e|}{m_e} = \frac{\hbar}{2m_e}. \tag{7}$$

In equation (7), m_e is the electron mass, and $|J_e| = \hbar/2$ is the electron spin.

In general, for a black hole with mass M, charge Q, and angular momentum J, the mass functions $m(r)$ for both classical and quantum Kerr (K) and Kerr–Newman (KN) metrics do not depend on the spin parameter $a = J/M$ and are therefore equal to the mass functions for the classical and quantum Schwarzschild and Reissner–Nordström metrics.

For the electron, the quantum mass function is

$$m^e_{KNq}(r) = m^e_{RNq}(r) = G m_e \frac{2}{\pi}Si\left(\frac{k^e_{UV}}{\hbar c}r\right) - \frac{G e^2}{2r}\left[1 - \cos\left(\frac{k^e_{UV}}{\hbar c}r\right)\right] = G m_e \frac{2}{\pi}Si\left(\frac{r}{R^e_S}\right) - \frac{G e^2}{2r}\left[1 - \cos\left(\frac{r}{R^e_S}\right)\right]. \tag{8}$$

where $Si(x) = \int_0^x \frac{\sin x'}{x'}dx'$ is the sine integral function. According to equation (2),

$$R^e_S = \frac{\pi}{8}\frac{e^2}{m_e c^2} = 1.11\cdot 10^{-13}\ \text{cm}. \tag{9}$$

According to equation (1), the maximum (cut-off) energy of gravitons is

$$k^e_{UV} = \frac{\hbar c}{R^e_S} = 178\ \text{MeV}.$$

[1] Below we will use units with the velocity of light $c = 1$. When calculating the numerical values of the theory parameters, we will use the value $c = 3\cdot 10^{10}$ cm/s.

The asymptotics of the quantum mass function (8) are

$$m^e_{KNq}\Big|_{r\to\infty} = G m_e, \tag{10}$$

$$m^e_{KNq}\Big|_{r\to 0} = \frac{1}{18}\frac{G m_e}{\pi}\left(\frac{r}{R^e_S}\right)^3 \to 0. \tag{11}$$

According to equation (10), the quantum KN metric becomes asymptotically flat as $r \to \infty$.

For the classical KN metric, the mass function $m^{cl}_{KN} = 0$ at $r_e = e^2 / 2m_e$, i.e. at $r = r_e$, the classical metric is flat in this limit [7]. For the quantum Kq and KNq metrics, the spacetime curvature persists throughout the entire interval $r \in (0,\infty)$ [5].

2.2 Modified Reissner–Nordström Geometry

The quantum RNq metric [4] can be obtained from equation (3) by setting $a_e = 0$:

$$ds^2_{RNq} = \left(1 - \frac{2m^e_{RNq}(r)}{r}\right)dt^2 -$$

$$-\frac{1}{1 - \dfrac{2m^e_{RNq}(r)}{r}}dr^2 - r^2\left(d\theta^2 + \sin^2\theta d\varphi^2\right), \tag{12}$$

where $m^e_{RNq}(r)$ is given in equation (8).

The quantum RNq metric is asymptotically flat as $r \to \infty$ (see equation (10)). The $g_{00} = -1 / g_{11}$ component at $r \to 0$ is

$$g_{00} = 1 - \frac{G m_e}{9\pi c^2 R^e_S}\left(\frac{r}{R^e_S}\right)^2 =$$

$$= 1 - 2.15\cdot 10^{-44}\left(\frac{r}{R^e_S}\right)^2, \tag{13}$$

meaning that the metric (12) becomes flat at $r = 0$.

2.3 Characteristics of electron models

Let's present some characteristic values for the electron:

$$m_e = 9.1\cdot 10^{-28}\ \text{g},\ e^2 = 2.31\cdot 10^{-19}\ \text{erg}\cdot\text{sm},$$

$$\text{spin}: \frac{\hbar}{2} = 0.5\cdot 1.054\cdot 10^{-27}\ \text{erg}\cdot\text{sm},$$

$$G = 6.67\cdot 10^{-8}\ \frac{\text{cm}^3}{\text{g}\cdot\text{s}^2},\ c = 3\cdot 10^{10}\ \frac{\text{sm}}{\text{s}},$$

$$R^e_H = \frac{2G m_e}{c^2} = 1.35\cdot 10^{-55}\ \text{cm},$$

$$\frac{Ge^2}{c^4} = \left(1.38\cdot 10^{-34}\right)^2\ \text{cm}^2,$$

$$a_e^2 = \left(\frac{\hbar}{2m_e c}\right)^2 = \left(1.93\cdot 10^{-11}\right)^2\ \text{cm}^2,$$

$$\beta_1 = \frac{Ge^2}{c^4}\frac{4}{\left(R^e_H\right)^2} = 4.2\cdot 10^{42},$$

$$\beta_2 = \frac{4a_e^2}{\left(R^e_H\right)^2} = 8.2\cdot 10^{88},\ \text{т.е. } \beta_1 + \beta_2 \gg 1,$$

$$R_{cl} = \frac{e^2}{m_e c^2} = 2.82\cdot 10^{-13}\ \text{cm},$$

$$R^e_S = \frac{\pi}{8}\frac{e^2}{m_e c^2} = 1.11\cdot 10^{-13}\ \text{cm},$$

$$k^e_{UV} = \frac{\hbar c}{R^e_S} = 178\ \text{MeV},$$

$$\frac{R^e_S}{R^e_H} = \frac{1.11\cdot 10^{-13}}{1.35\cdot 10^{-55}} = 0.82\cdot 10^{42}.$$

We see that for the electron, $\beta_1 + \beta_2 \gg 1$, $R^e_S / R^e_H \gg 1$. This means that in the models of electron with the RNq and KNq quantum metrics, the event horizons are absent [8]. The proposed electron models represent either rotating (KNq) or non-rotating (RNq) collapsars without event horizons and with quantum cores defined by coherent states of gravitons with a maximum energy of $k^e_{UV} = 178$ Me V.

2.4 Electromagnetic potentials

For the classical Reissner–Nordström and Kerr–Newman metrics with mass M and charge Q, the mass function consists of two terms:

$$m^{cl}(r) = \left(m^{cl}(r)\right)_M + \left(m^{cl}(r)\right)_Q = GM - \frac{GQ^2}{2r}. \tag{14}$$

The "charge" part of the mass function

$$\left(m^{cl}(r)\right)_Q = -GQ^2/2r$$

ensures that the "charge" components of the Einstein tensor, divided by $8\pi G$, match the corresponding components of the electromagnetic field energy-momentum tensor derived from Maxwell's equations:

$$\frac{\left(G_{\mu}{}^{\nu}\right)_{Q}}{8\pi G} = \left(T_{\mu}{}^{\nu}\right)_{em}.$$

For the classical KN geometry, the electromagnetic potentials A_{μ} are chosen as follows [9]:

$$A_{\mu} = \frac{Qr}{\rho^2}\left(1,0,0,-a\sin^2\theta\right). \quad (15)$$

Electromagnetic fields at $r \to \infty$ manifest as a superposition of the Coulomb field and the magnetic dipole field $\mu = Qa$. The gyromagnetic ratio $\mu/|J| = Q/m$, which coincides with the gyromagnetic ratio for a Dirac electron. The complex internal electromagnetic structure of the classical KN metric source is discussed, for example, in [10].

For the classical Reissner–Nordström (RN) metric, when $(a = 0)$ in equation (15), only the scalar Coulomb potential remains $A_0 = Q/r$.

For the regular quantum electron metrics (considering the relation between m_e and e^2 from equation (9)), the "charge" part of the mass function can be retained as in the classical RN and KN metrics. In this case, the mass function (8) becomes:

$$m^e_{RNq}(r) = m^e_{KNq}(r) =$$

$$= G m_e\left[\frac{2}{\pi}\mathrm{Si}\left(\frac{r}{R^e_S}\right) + \frac{4}{\pi}\frac{\cos\left(r/R^e_S\right)}{r/R^e_S}\right] - \frac{Ge^2}{2r}. \quad (16)$$

Thus, the electromagnetic properties of the proposed electron models coincide with the electromagnetic properties of the sources of the classical Reissner–Nordström and Kerr–Newman metrics.

2.5 Electron's self-energy

In the study [5], we established that for

$$R_S = R^{reg}_S = \pi Q^2/8M$$

the energy of a rotating charged quantum black hole equals $E = M$ (see also the Appendix). A similar equality holds for the RNq quantum metric at any value of R_S. For electron models in natural units:

$$R^e_S = \pi e^2/8m_e c^2 = 1.11\cdot 10^{-13}\,\mathrm{cm}.$$

The equality $E = m_e$ means that the electron's self-energy E_{em} is zero.

3. DISCUSSION

We have examined two quantum models of the electron based on modified Reissner–Nordström (RNq) and Kerr–Newman (KNq) metrics. Can we currently favor one model over the other? To answer this question, let us compare some characteristics of the considered models under the condition

$$R_S = R^e_S = \frac{\pi e^2}{8m_e}.$$

Table: Comparison of electron model characteristics in Reissner–Nordström (RNq) and Kerr–Newman (KNq) quantum geometries

	Electron model characteristic	RNq	KNq
1	$E_e = m_e$, $E_{em} = 0$	+	+
2	Weak energy condition	+	–
3	$\lvert J\rvert = \frac{\hbar}{2}$, Dirac gyromagnetic ratio $\frac{\mu}{\lvert J\rvert} = \frac{e}{m_e}$	–	+
4	Absence of event horizons	+	+
5	Finiteness of the GRT quantities, such as the mass function, Ricci tensor, Kretschmann scalar, etc.	+	+
6	Compatibility with the Maxwell equations	+	+
7	Stationary bound states in the fields of regular black holes	+	–

In the table, the symbols "+" and "–" indicate the presence or absence of key characteristics in the considered models.

Let us briefly discuss points 1–7 of the table.

Point 1. For both models:

$$E_e = m_e c^2, E_{em} = 0.$$

We found an important aspect: gravity in the charged quantum Kerr–Newman (rotating) and Reissner–Nordström (non-rotating) metrics with $R_S = R^e_S$ compensates for the electromagnetic component in the expressions for the total energy of the quantum black hole.

In classical electrodynamics, the self-energy of a charged particle $E^{cl}_{em} = e^2/2r$ diverges linearly as $r \to 0$. In quantum field theory, the self-energy of a charged particle is determined by an infinite series in perturbation theory with logarithmic divergence terms.

Point 2. For the RNq quantum geometry, the energy density $\rho_{\varepsilon}(r)$, radial pressure $p_1(r)$, and stresses $p_2(r) = p_3(r)$ take the following form [4]:

$$\rho_{\varepsilon}(r) = -p_1(r) = \frac{m_e}{\pi^2\left(R^e_S\right)^3}\times$$

$$\times\left[\frac{1}{\left(r/R_S^e\right)^4}\left(1-\cos\left(\frac{r}{R_S^e}\right)\right)-\right.$$

$$\left.-\frac{1}{2\left(r/R_S^e\right)^3}\sin\left(\frac{r}{R_S^e}\right)\right], \quad (17)$$

$$p_2(r)=p_3(r)=\frac{m_e}{\pi^2\left(R_S^e\right)^3}\times$$

$$\times\left[\frac{1}{\left(r/R_S^e\right)^4}\left(1-\cos\left(\frac{r}{R_S^e}\right)\right)+\frac{1}{4\left(r/R_S^e\right)^2}\cos\left(\frac{r}{R_S^e}\right)-\right.$$

$$\left.-\frac{3}{4\left(r/R_S^e\right)^3}\sin\left(\frac{r}{R_S^e}\right)\right]. \quad (18)$$

At $r\to 0$, we have $\rho_\varepsilon(r)\to K/24$, $p_i(r)\to -K/24$, where $i=1,2,3$ and $K=m_e\Big/\pi^2\left(R_S^e\right)^3$. Thus, for the RNq quantum geometry near $r=0$, the weak energy condition $\rho_\varepsilon\geq 0$, $\rho_\varepsilon+p_i\geq 0$, $i=1,2,3$ is satisfied.

Specifically, equations (17) and (18) show that at $r=0$ $\rho_\varepsilon=K/24$, $\rho_\varepsilon+p_i=0$, $i=1,2,3$.

For the RNq quantum geometry at $r=0$, the energy dominance condition $\rho_\varepsilon\geq|p_i|$, $i=1,2,3$ also holds. In our case: $\rho_\varepsilon=|p_i|$.

For the Kerr–Newman quantum geometry, the asymptotics of the energy density $\rho_\varepsilon(r,\mu)$ at $r\to 0$ follow from equation (7) in [5] (here and below, $\mu=\cos\theta$):

$$\rho_\varepsilon(r,\mu)=\frac{K}{12}\frac{\mu^2-1}{\mu^4}\left(\frac{r}{a_e}\right)^2,\ \mu\neq 0,$$
$$\rho_\varepsilon(r,\mu)=84K,\ \mu=0. \quad (19)$$

At $\mu\neq 0,\pm 1$ the energy density near $r=0$ is negative. In this case, none of the energy conditions are satisfied.

Point 3. In the KNq quantum model, it is possible to introduce the spin modulus $|J|=\hbar/2$, satisfying the Dirac gyromagnetic ratio. However, introducing the quantum spin operator $\mathrm{S}=(\hbar/2)\sigma$ is complicated when the classical definition of angular momentum is used in the Kerr–Newman geometry. Above,, σ_i are two-dimensional Pauli matrices.

In the RNq quantum geometry, the angular momentum J is zero. In the RNq electron quantum model, the spin operator S and the gyromagnetic ratio e/m_e are pure quantum properties defined externally.

Point 4. In both these models, event horizons are absent.

Point 5. In both models, general relativity (GR) quantities such as the mass function, Ricci tensor, Kretschmann scalar, and others remain finite.

Point 6. The RNq and KNq quantum geometries are consistent with Maxwell's equations (see Section 2.4 of this study). However, the electromagnetic structure of the RNq model is significantly simpler than that of the KNq model. In the RNq quantum model, the source of the electromagnetic field is a point electric charge e located at the system's center $(r=0)$. At large distances, the electromagnetic field behaves as a Coulomb field.

In contrast, the source of the electromagnetic field in the KNq quantum model is a system of surface currents and electric charges distributed over a disk of radius $a_e=|J_e|/m_e c$ with the center at $r=0$ [10]. For $r\to\infty$, the electromagnetic field is a superposition of the Coulomb field and a magnetic dipole $\mu=ea$.

Point 7. In the RNq quantum geometry, the metric (12) becomes asymptotically flat as $r\to\infty$. Importantly, for both $R_S=R_S^e$ and $r\to 0$, the metric (12) is also flat (see Equation (13)). In this case, the problem of determining the eigenfunctions and eigenvalues of the Dirac equation for motion of fermions in the RNq fields can be solved by using single-valued boundary conditions from the analogous problem for the fermion motion in the Coulomb field in flat Minkowski space.

In the Kerr–Newman quantum geometry, the situation is different. At $r\to 0$ и $R_S=R_S^e$, the metric (3) remains non-flat and takes the following form:

$$ds_{KNq}^2=dt^2-\cos^2\theta dr^2-a_e^2\cos^2\theta d\theta^2-$$
$$-a_e^2\sin^2\theta d\varphi^2. \quad (20)$$

In [11, 12], it was shown that in this case, the Dirac equation has two quadratically integrable solutions, making it impossible to formulate a well-defined eigenvalue problem for fermions in the classical or quantum KN spacetime.

To establish a well-defined quantum mechanical problem, one must perform a self-adjoint extension of the Hamiltonian, which usually results in new boundary conditions near $r = 0$ (see, for example, [13, 14]).

4. CONCLUSION

We proposed two quantum electron models with zero self-energy based on the Reissner–Nordström [4] and Kerr–Newman [5] quantum geometries. A critical parameter for regularizing key GR quantities is the choice of $R_S^e = \pi e^2/8m_e c^2 \simeq 1.11\cdot 10^{-13}$ cm, where the cut-off energy of gravitons $k_{UV}^e = \hbar c/R_S^e \approx 178$ Me V.

The proposed models solve the long-standing problem of linear divergence in the self-energy of a charged particle in classical electrodynamics. In the considered models, gravity compensates for the electromagnetic component in the total energy expressions for the electron.

It can be hypothesized that with more advanced quantum gravity theories, the problem of infinite self-energy of charged fermions in quantum field theory will be resolved similarly.

Notably, when using the RNq quantum electron model, all classical and quantum electrodynamics effects can be calculated within the standard paradigm of an elementary particle with point mass m_e and electric charge $e < 0$. This is due to the extremely small values of the parameters $Gm_e/c^2 \simeq 0.7\cdot 10^{-55}$ cm and $Ge^2/c^4 \simeq 1.9\cdot 10^{-68}$ cm^2 in Equation (16) for the mass function $m_{RNq}^e(r)$.

As a result of neglecting the coefficients $\frac{Gm_e}{c^2}$ and $\frac{Ge^2}{c^4}$ the RNq geometry becomes the flat Minkowski space-time. In this case, we return to the domain of classical and quantum electrodynamics for charged leptons within the Standard Model.

FUNDING

This study was conducted within the framework of the scientific program of the National Center for Physics and Mathematics, section N5 “Particle physics and cosmology. Stage 2023-2025”.

ACKNOWLEDGMENTS

The authors thank S.Yu. Sedov for numerous discussions on some sections of this paper and A.L. Novoselova for significant technical assistance in preparing the article.

APPENDIX: ENERGY OF A CHARGED ROTATING BLACK HOLE WITH A QUANTUM CORE [5]

For the KN quantum metric, the total energy, defined by the volume integral of the energy density $T_0^{\,0} \equiv \rho_\varepsilon(r,\theta)$, is given by:

$$E = \int T_0^{\,0}\sqrt{-g}\,dV = \frac{1}{4G}\int_0^\infty dr\int_{-1}^1 d\propto\left(r^2 + a^2\propto^2\right)G_0^{\,0}(r,\propto) =$$

$$= \frac{1}{4G}\int_0^\infty dr\int_{-1}^1 d\mu\left[2\frac{r^4 + \left(\rho^2 - r^2\right)^2 + a^2\left(2r^2 - \rho^2\right)}{\rho^4}m'_{KN_q} - \frac{ra^2\left(1-\mu^2\right)}{\rho^2}m''_{KN_q}\right] =$$

$$= \frac{1}{4G}\int_0^\infty dr\left\{\left[8 - 4\frac{r}{a}\mathrm{arctg}\frac{a}{r}\right]\left[GM\frac{2}{\pi}\frac{\sin(r/R_S)}{r} + \frac{CQ^2}{2r^2}\left(1 - \cos\left(\frac{r}{R_S}\right)\right) - \frac{CQ^2}{2rR_S}\sin\left(\frac{r}{R_S}\right)\right] +\right.$$

$$+\left[2r - 2\frac{r^2}{a}\mathrm{arctg}\frac{a}{r} - 2a\,\mathrm{arctg}\frac{a}{r}\right]\left[GM\frac{2}{\pi}\frac{\cos(r/R_S)}{r/R_S}\frac{1}{R_S^2} - GM\frac{2}{\pi}\frac{\sin(r/R_S)}{(r/R_S)^2}\frac{1}{R_S^2} - \frac{CQ^2}{r^3}\left(1 - \cos\left(\frac{r}{R_S}\right)\right) +\right.$$

$$\left.\left. + \frac{CQ^2}{r^2R_S}\sin\left(\frac{r}{R_S}\right) - \frac{CQ^2}{2rR_S^2}\cos\left(\frac{r}{R_S}\right)\right]\right\} = M + \frac{1}{2}\frac{|J|}{R_S} - \frac{\pi}{16}\frac{Q^2|J|}{M}\frac{1}{R_S^2} = M + \frac{1}{2}\frac{|J|}{\hbar}k_{UV} - \frac{\pi}{16}\frac{Q^2|J|}{M}\frac{k_{UV}^2}{\hbar^2}.$$

For the K and KN metrics:

$$\sqrt{-g} = \rho^2 \sin\theta; \rho^2 = r^2 + a^2\mu^2, \mu = \cos\theta;$$

$$m'_{KNq} \equiv \frac{dm_{KNq}}{dr}, m''_{KNq} \equiv \frac{d^2 m_{KNq}}{dr^2},$$

$$m_{KNq} = G M \frac{2}{\pi} Si\left(\frac{r}{R_S}\right) - \frac{G Q^2}{2r}\left(1 - \cos\left(\frac{r}{R_S}\right)\right).$$

When the condition

$$R_S = R_S^{reg} = \frac{\pi}{8}\frac{Q^2}{Mc^2}$$

is satisfied, the total energy of the quantum charged rotating collapsar equals zero: $E = Mc^2$.

Under this condition, the key general relativity (GR) quantities, such as the mass function $m(r)$, the Ricci tensor $R_{\mu\nu}(r,\theta)$, and the Kretschmann scalar $K(r,\theta)$, become regular and finite throughout the entire spacetime.

REFERENCES

1. *C. V. Costa, D. M. Gitman, and A. E. Shabad*, Phys. Scr. **90**, 074012 (2015).
2. *T. C. Adorno, D. M. Gitman, A. E. Shabad, A. A. Shishmarev*, Izv. Vuzov **59**, 45 (2016) [in Russian].
3. *R. Casadio*, Int. J. Mod. Phys. D **31**, 2250128 (2022); arxiv: 2103.00183v4 (gr-qc).
4. *R. Casadio, A. Giusti, and J. Ovalle*, Phys. Rev. D **105**, 124026 (2022); arxiv: 2203.03252v2 (gr-qc).
5. *V. P. Neznamov, S. Yu. Sedov, and V. E. Shemarulin*, Int. J. Mod. Phys. A **39**, 2450012 (2024).
6. *M. Cürses and F. Cürsey*, J. Math. Phys. **16**, 2385(1975).
7. *C. A. Lopez*, Phys. Rev. D **30**, 313 (1984).
8. *V. P. Neznamov, I. I. Safronov, V. E. Shemarulin*, Doklady Physics **68**, No. 7, 209 (2023).
9. *B. Carter*, Phys. Rev. **174**, 1559 (1968).
10. *C. L. Pekeris, K. Frankowski*, Phys. Rev. A **36**, 5118 (1987).
11. *C. L. Pekeris*, Phys. Rev. A **35**, 14 (1987).
12. *C. L. Pekeris, K. Frankowski*, Phys. Rev. A **39**, 518 (1989).
13. *D. M. Gitman, I. V. Tyutin, and B. L. Voronov*, Self-Adjoint Extensions in Quantum Mechanics: General Theory and Applications to Schrödinger and Dirac Equations with Singular Potentials, Progr. Math. Phys. **62**, New York: Birkhäuser, 2012).
14. *A. E. Shabad*, J. Phys. A **38**, 7419 (2005); arXiv: hep-th/0502139 (2005).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 64–75

SOLIDS AND LIQUIDS

NEUTRON CHANNELING IN A MAGNETIC PLANAR WAVEGUIDE

© 2025 S. V. Kozhevnikov[a,*], Yu. N. Khaydukov[b, c, d]

[a]Frank Laboratory of Neutron Physics, Joint Institute for Nuclear Research, Dubna, Moscow region, Russia
[b]Skobeltsyn Institute of Nuclear Physics Lomonosov Moscow State University, Moscow, Russia
[c]Max Planck Institut für Festkörperforschung, Stuttgart, Germany
[d]Max Planck Society Outstation at FRM–II, Garching, Germany
*e-mail: kozhevn@nf.jinr.ru

Received May 08, 2024

Reveised June 03, 2024

Accepted June 10, 2024

Abstract. Tri-layered waveguide transforms a conventional collimated neutron beam into a narrow divergent microbeam. Propagation of neutrons in a waveguide with enveloping magnetic layers is investigated. Intensity of the neutron microbeam emitted from the end face of the nonmagnetic middle layer is registered. Neutron channeling length is defined experimentally in dependence on the sign of polarization of the incident beam.

Keywords: *layered nanostructures, neutron resonators, waveguides, neutron channeling, neutron microbeam*

DOI: 10.31857/S00444510250106e6

1. INTRODUCTION

Neutron scattering is a powerful non-destructive method for studying magnetic structures, polymers, and biological objects due to the unique properties of neutrons: the presence of an intrinsic magnetic moment, high penetrating ability, and isotopic sensitivity. The properties of neutron and X-ray radiation differ significantly, making them complementary methods. For example, polarized neutron beams are a unique tool for studying magnetic materials within the bulk of a substance, which is inaccessible to X-rays due to their low penetrating ability.

The width of the neutron beam determines the spatial resolution and the scale of the studied objects. The typical beam width in neutron experiments ranges from 0.1 to 10 mm. To study local microstructures on the scale of tens of micrometers, very narrow neutron beams are required. For this purpose, various focusing devices are being developed (parabolic mirror neutron guides, refractive lenses, curved monochromator crystals, etc.) [1], capable of compressing the neutron beam to 50 μm. Achieving a smaller beam width is hindered by limitations imposed by the physical properties of the materials used and the technology of their processing. Another problem with these devices is their inability to effectively isolate a "pure" microbeam. For example, parabolic mirror neutron guides form a highly structured beam in space, refractive lenses focus only 20–30% of the initial beam, and capillary lenses generate significant background noise.

In [2], the profile of a microbeam after passing through an aperture formed by neutron-absorbing crystal blades $Gd_2Ga_5O_{12}$ (or GGG) was calculated. The resulting microbeam had a central part about 100 μm wide and wings ranging from 10 to 20 μm. The study also demonstrated a method for obtaining a microbeam through total reflection of neutrons from a silicon substrate. This method has undeniable advantages: high intensity (~1000 neutrons/s), low background (~2 neutrons/min), and compatibility with time-of-flight techniques. However, the practically achievable microbeam width at a neutron wavelength of 4.0 Å and an 8 mm wide silicon substrate still remains around 30 μm.

The record holders for the minimum width of neutron microbeams are triple-layer waveguides (Fig. 1). Their operating principle is as follows. A collimated neutron beam with an angular divergence $\delta\alpha_i$ falls in a vacuum (medium 0) onto the surface of the waveguide at a small grazing

angle α_i. The neutrons then tunnel through a thin upper layer (medium 1) with a thickness of $a = 5-20$ nm and enter the middle layer (medium 2) with d=100–200 nm. They are almost completely reflected from the relatively thick lower layer (medium 3), deposited on a thick substrate (e.g., glass). Some of the neutrons tunnel back through the upper layer and exit the waveguide as a reflected beam $\alpha_f = \alpha_i$. Another portion of the neutrons reflects from the upper thin layer 1 and returns to the middle layer 2. As a result of multiple reflections, the neutrons propagate along the middle layer as if through a channel and exit from its end as a microbeam with an angular divergence $\delta\alpha_f$. The main contribution to the angular divergence $\delta\alpha_f$ of the microbeam comes from Fraunhofer diffraction $\delta\alpha_F$ at the narrow slit d formed by the waveguide channel $\delta\alpha_F \propto \lambda / d$, where λ is the neutron wavelength.

Layered neutron waveguides have been well studied to date. In [3], an unpolarized neutron microbeam was obtained from the end of a triple-layer waveguide, while in [4], a polarized beam was achieved. The contribution of Fraunhofer diffraction $\delta\alpha_F$ to the angular divergence of the neutron microbeam was experimentally determined in [5, 6, 7]. In [2, 8], a polarized neutron microbeam from a waveguide was used for spatial scanning of a 190 μm diameter microwire made of amorphous magnetic material. At a distance of 1 mm from the waveguide exit, with a neutron wavelength of 4.0 Å, a waveguide channel width of 150 nm, and an angular divergence 0.15°, the calculated microbeam width at the sample location was 2.6 μm. With a microbeam intensity of approximately 1 neutron/s, statistically significant data were obtained within about 10 hours. The experimental setup is described in detail in [2]. The advantages of planar waveguides include the record-low width of the neutron microbeam and a relatively simple method for separating the microbeam from the background. Their obvious disadvantages are low intensity and relatively high beam divergence. However, the commissioning of more powerful neutron sources (SNS, ESS, PIK, IBR-3) may make the use of layered waveguides more accessible.

In planar waveguides, two phenomena are observed simultaneously – resonant enhancement of neutron standing waves and neutron channeling. The theory of neutron resonances in layered waveguides is described in [9]. Let us introduce the following notations:

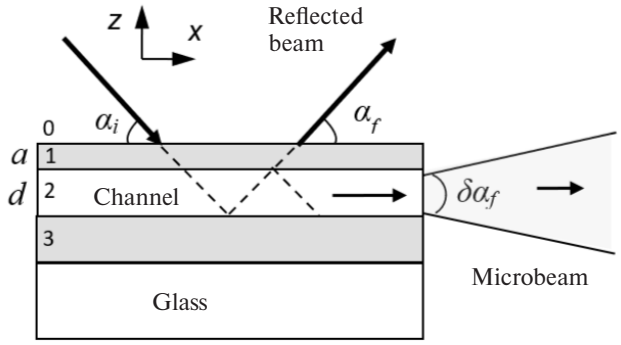

Fig. 1. Principle of operation of a planar neutron waveguide

$$k_{0z} = \frac{2\pi}{\lambda}\sin\alpha_i,$$

$$k_{1z} = \sqrt{k_{0z}^2 - \rho_1}, \quad k_{2z} = \sqrt{k_{0z}^2 - \rho_2},$$

$$k_{0x} = \frac{2\pi}{\lambda}\cos\alpha_i.$$

Here, ρ_1 is the neutron scattering length density (SLD) for the upper layer 1, and ρ_2 is the SLD for the waveguide layer 2. The general form of the neutron wave function is given by:

$$\Psi(k_{0z}, z) = A\exp(ik_{0z}z),$$

where A is the amplitude of the wave function. Then, we obtain the condition $|\Psi|^2 = |A|^2$. Inside the middle layer, the wave function takes the form:

$$\Psi(z) = A\left[\exp(-ik_{2z}z) + R_{23}\exp(ik_{2z}z)\right],$$

where R_{23} is the amplitude of the reflected neutron wave function from the lower layer 3. The amplitude A is determined from the self-consistent equation for the neutron wave function in layer 2, if the origin z=0 is aligned with the boundary between layers 1 and 2:

$$A = T_{02}\exp(ik_{2z}d) + R_{21}R_{23}\exp(ik_{2z}2d)A, \quad (1)$$

where T_{02} is the amplitude of the transmitted neutron wave function from vacuum (medium 0) into medium 2, and R_{21} is the amplitude of the reflected neutron wave function in medium 2 from layer 1. From the self-consistent equation (1), we find:

$$|\Psi|^2 = |A|^2 = \frac{|T_{02}|}{|1 - R_{21}R_{23}\exp(2ik_{2z}d)|}. \quad (2)$$

The quantity $|A|^2$ in equation (2) exhibits resonant maxima under the periodic conditions for the phase of the neutron wave function:

$$\Phi(k_{0z}) = 2k_{2z}d + \arg(R_{21}) + \arg(R_{23}) = 2\pi n, \quad (3)$$

where $n = 0,1,2,\ldots$ is the resonance order. If the neutron wavelength is fixed, the grazing angle of the incident beam has resonances depending on the angle α_{in}. If the time-of-flight method is used, the grazing angle is fixed, and the final neutron spectrum exhibits resonances depending on the wavelength λ_n. In [10], it was experimentally shown using a time-of-flight reflectometer that the spectral width of neutron resonances increases with the divergence $\delta\alpha_i$ of the incident beam.

The parameter $|A|^2$ represents the neutron density enhancement coefficient inside the middle layer, and for various resonators, it can reach values of the order of 10^1–10^3. Layered resonators are used to amplify the weak interaction of neutrons with matter [11]. Neutron resonances appear as weak minima in the coefficients of neutron specular reflection and as corresponding resonance-enhanced maxima of secondary characteristic radiation or specific neutron scattering. During neutron interactions with certain elements and isotopes, secondary characteristic radiation, such as gamma rays [12] and alpha particles [13], is generated due to nuclear reactions. The experimental setup and the method of neutron reflectometry with the detection of secondary radiation are described in detail in [14–17].

Specific neutron scattering within the resonator can also include neutrons that experience spin-flip interactions with magnetically non-collinear layered structures [18, 19, 20], incoherent scattering from hydrogen [21], and off-specular scattering from interlayer roughness [22, 23] and domain structures [24, 25]. The high sensitivity of neutron resonance positions to changes in the SLD of the resonant layer has been utilized to detect small variations in hydrogen concentration within the resonator [26, 27]. These resonators can be applied as sensors in hydrogen storage systems.

Another type of specific neutron scattering is neutron channeling. A neutron beam propagating along the middle layer can exit through the waveguide surface as a collimated beam of standard width or from the channel end as a narrow divergent microbeam (see Fig. 1). The intensities of both neutron beams exhibit resonance maxima depending on energy. In [28], the idea of using planar neutron waveguides to determine weak magnetization of films on the order of 10^2 G was proposed. This idea was experimentally implemented in [29] and [30]. In the three-layer waveguide, the outer layers were non-magnetic, while the investigated ferrimagnetic films $TbCo_5$ [29] and $TbCo_{11}$ [30] acted as the middle waveguide layer. The magnetization value is determined directly from the difference in the resonance positions, which varies by about n=0 for the incident beam polarizations "+" and "–". Moreover, registering the microbeam allows effective separation of the useful signal from the background, originating from the specularly reflected, refracted, and bypassing beams. In this study, we examine a waveguide where the outer layers are magnetic and the middle layer is non-magnetic (see Fig. 2). In such waveguides, the neutron density enhancement coefficient within the waveguide channel depends on the neutron spin projection "+" or "–" relative to the magnetization vector direction. In [31], the idea was proposed to control the chain reaction of uranium fission within the non-magnetic waveguide layer by magnetizing the outer layers using an applied magnetic field. This approach alters the parameter x_e, which characterizes the exponential attenuation of the neutron density, known as the channeling length.

In [32], it was theoretically demonstrated that during neutron propagation along the waveguide channel, the neutron wavefield attenuates as $\exp(-x / x_e)$, where x is the distance under the unilluminated surface of the waveguide. The expression for the neutron channeling length was derived as:

$$x_e = \frac{k_x d}{k_{2z}\left|\ln\left|R_{21}R_{23}\right|\right|}. \quad (4)$$

If the lower layer is sufficiently thick, we can assume $R_{23} = 1$. If the neutron reflection amplitude from the upper layer is close to unity $R_{21} \approx 1$, the neutron transmission coefficient through the upper layer becomes a small parameter:

$$T = |T_{20}| = 1 - |R_{21}|$$

Thus, we can write an approximate expression:

$$\left|\ln\left|R_{21}R_{23}\right|\right| \approx \left|\ln(1-T)\right| \approx T.$$

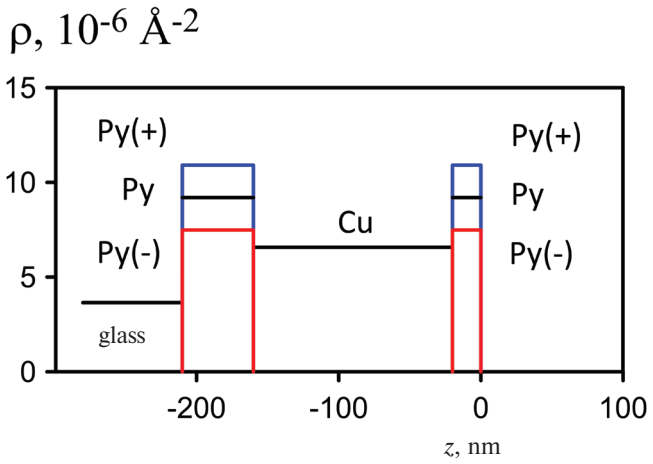

Fig. 2. Neutron scattering length density (SLD) of the waveguide with magnetic outer layers as a function of the coordinate z perpendicular to the layers. The notation Py(+) corresponds to the polarization "+" and a film magnetization of 7.2 kG, Py (—) corresponds to the polarization "–" and a film magnetization of 7.2 kG, while Py represents the unpolarized beam and the demagnetized sample.

In this case, we obtain a simplified expression for the neutron channeling length:

$$x_e \approx \frac{k_x d}{k_{2z} T}. \qquad (5)$$

The neutron channeling length can be determined experimentally. A strip of neutron-absorbing material is applied near the exit edge of the waveguide, creating an unilluminated region of length x. Then, by varying the absorber's position and changing the length of the unilluminated region x, the intensity $I(x)$ of the microbeam emerging from the waveguide end is measured. For normalization, the microbeam intensity without the absorber $I(x = 0)$ is recorded.

According to the channeling theory [32], the intensity of the neutron microbeam from the waveguide channel end attenuates exponentially with increasing length of the unilluminated waveguide surface:

$$I(x) / I(x = 0) = \exp(-x / x_e). \qquad (6)$$

From the experimental dependence of the microbeam intensity (6), the neutron channeling length x_e can be determined. For various waveguides, this value typically ranges between 0.5 and 5.0 mm.

Various materials are used as neutron absorbers: Gd_2O_3 powder, Cd plates, or boral (an aluminum-boron carbide composite) bars. Fig. 3 shows the

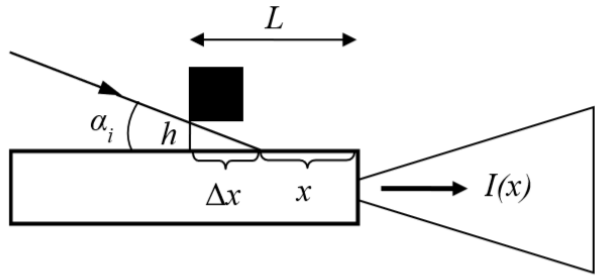

Fig. 3. Experimental setup for determining the neutron channeling length using a sliding absorber bar.

experimental setup with a sliding boral block. Due to the block's curvature, an air gap of approximately with $h \approx 10\mu m$ forms between the block and the waveguide surface. As a result, part of the waveguide surface under the absorber, with a length Δx of about 1.5 mm, remains illuminated by the incident neutron beam. In the experiment, the intensity $I(L)$ of the neutron microbeam is recorded as a function of the distance L from the waveguide exit edge to the front edge of the absorber. The coordinate $L = \Delta x + x$ includes both the illuminated length Δx and the unilluminated length x of the waveguide surface under the absorber. By transforming the coordinates, the dependence of the microbeam intensity on the unilluminated surface length $x = L - \Delta x$ is determined. The intensity $I(x)$ of the microbeam with the fully illuminated waveguide surface $I(x = 0)$ is used for normalization. The value Δx not need to be known in advance; it is determined automatically during the data processing, described in detail in Section 3.

The experimental setup and various methods for measuring the neutron channeling length are presented in [33]. Two neutron absorbers were compared: a sliding boral block and Gd_2O_3 powder. The advantage of the powder lies in its low background in the microbeam and simpler data processing, as there is no air gap between the surface and the powder. However, the drawbacks of the powder include: 1) significant time consumption when changing the absorber width; 2) practical infeasibility for neutron channeling lengths shorter than 1 mm;

The advantages of the sliding boral block are: 1) precise position control using a micrometer screw; 2) faster repositioning compared to the powder absorber; 3) suitability for determining short channeling lengths less than 1 mm. The drawbacks of the sliding boral block are higher background levels

in the microbeam compared to the powder absorber and more complex data processing.

The same study [33] experimentally demonstrated that the exponential attenuation parameter of neutron density in the reflection geometry is smaller than the channeling length in the microbeam geometry from the waveguide end.

The phenomenon of neutron channeling in three-layer waveguides was first observed in the reflection geometry in [34]. The first experimental measurement of the neutron channeling length in the microbeam geometry was performed in [35] using an absorbing powder on the surface. In [36], experiments were conducted with a Cd plate on the sample surface. A comprehensive review of studies on planar neutron waveguides is provided in [37], showing that the channeling length (5) depends on the resonance order $n = 0,1,2\ldots$ and the waveguide parameters – upper layer thickness a, channel width d, and the depth of the potential well defined by the scattering length density (SLD) contrast $\Delta\rho = \rho_1 - \rho_2$.

The following relationships were derived for the resonance order $n = 0$:

$$\ln x_e \propto a, \quad \ln x_e \propto d, \quad \ln x_e \propto \Delta\rho$$

and the first three resonance orders $n = 0, 1, 2$:

$$x_e \propto 1/(n+1)$$

Experiments with the sliding boral block determined the neutron channeling length as a function of resonance order and upper layer thickness [38], waveguide channel width [39], and potential well depth for various waveguides [40]. The experimental results confirmed theoretical predictions.

In this study, we experimentally determine the neutron channeling length in a waveguide with magnetic outer layers, where the potential well depth varies depending on the neutron beam polarization.

2. CALCULATIONS

Calculations were performed for the Py(20 nm)/Cu(140 nm)/Py(50 nm)//glass waveguide. Permalloy (Py) is a magnetic Fe(20.6 at.%)Ni(79.4 at.%) alloy with a narrow hysteresis loop. Figure 2 shows the neutron scattering length density (SLD) profile of the waveguide as a function of the coordinate z perpendicular to the layers. The designations Py(+) and Py(–) correspond to the SLD of saturated permalloy for neutron spins "UP" and "DO," respectively, while Py represents the SLD for the fully demagnetized state of permalloy. As seen in the figure, the SLD of permalloy changes depending on the neutron spin direction. The permalloy magnetization used for calculations is 7.2 kG, and the neutron wavelength is 4.26 Å. Fig. 4 presents the calculated squared modulus of the neutron wavefunction $|\Psi|^2$ as a function of the incident beam's grazing angle α_i and the coordinate z perpendicular to the layers. Fig. 4a shows "+" polarization, Fig. 4b shows calculations for the unpolarized NM beam, while Fig. 4c shows "–" polarization. Resonances of orders $n = 0,1,2\ldots$,... are visible, with the most intense ones located within the total reflection region below the horizontal dashed line. The neutron density enhancement coefficient reaches 30 for the UP polarization and the $n = 0$ resonance. Notably, the two-dimensional neutron density maps differ depending on the neutron beam polarization. As the waveguide potential well depth decreases, the resonance positions shift to lower incident angles, the distance between resonances decreases, and the resonance peak intensities also decrease.

Fig. 5a shows the neutron specular reflection coefficients for the UP polarization (thin line), the unpolarized NM beam (dashed line), and the DO polarization (thick line) as a function of the incident beam's grazing angle. It is evident that the total reflection region shifts toward smaller grazing angles for the NM and DO polarizations compared to the UP polarization. Additionally, minima in the reflection coefficients appear in the total reflection region, corresponding to resonance conditions $n = 0,1,2\ldots$.

In Fig. 5b, the square modulus of the neutron wavefunction $|\Psi|^2$ (in relative units), integrated over the coordinate z within the waveguide channel, is shown as a function of the incident beam's grazing angle. Resonance peaks $n = 0,1,2\ldots$ corresponding to the resonance order are clearly visible. If the peak value for the $n = 0$ resonance with UP polarization is normalized to 1.0, the corresponding peak for the unpolarized beam is 0.8, while for DO polarization it is 0.4. Thus, the square modulus of the neutron wavefunction $|\Psi|^2$ depends significantly on the potential well depth of the waveguide.

Fig. 6 shows the neutron channeling length $n = 0$ resonance as a function of the waveguide's potential

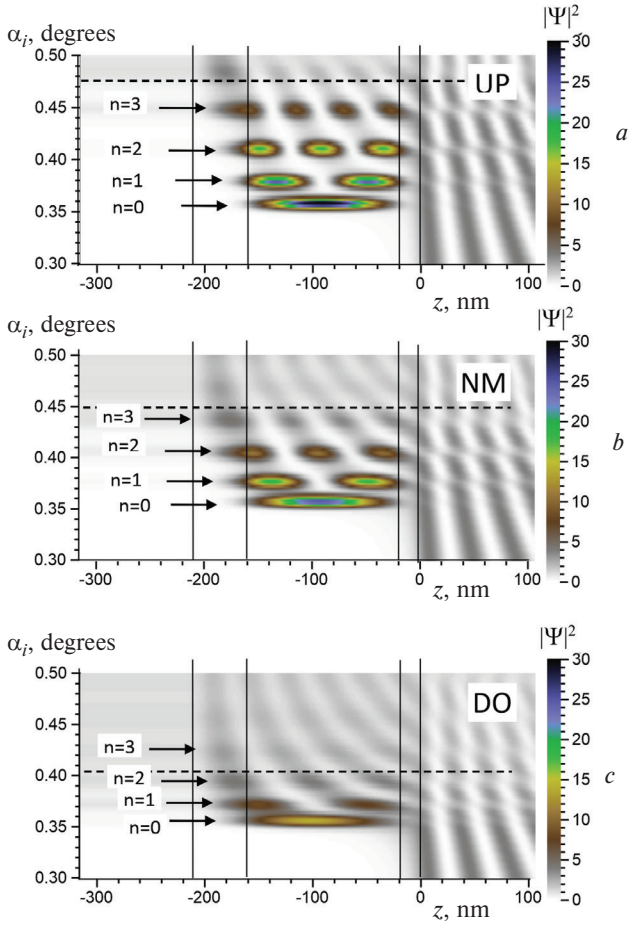

Fig. 4. Calculated squared modulus of the neutron wavefunction as a function of the grazing angle of the incident beam and the coordinate perpendicular to the layers for different initial beam polarizations: (a) UP; (b) Unpolarized NM beam; (c) DO. Neutron wavelength: 4.26 Å.

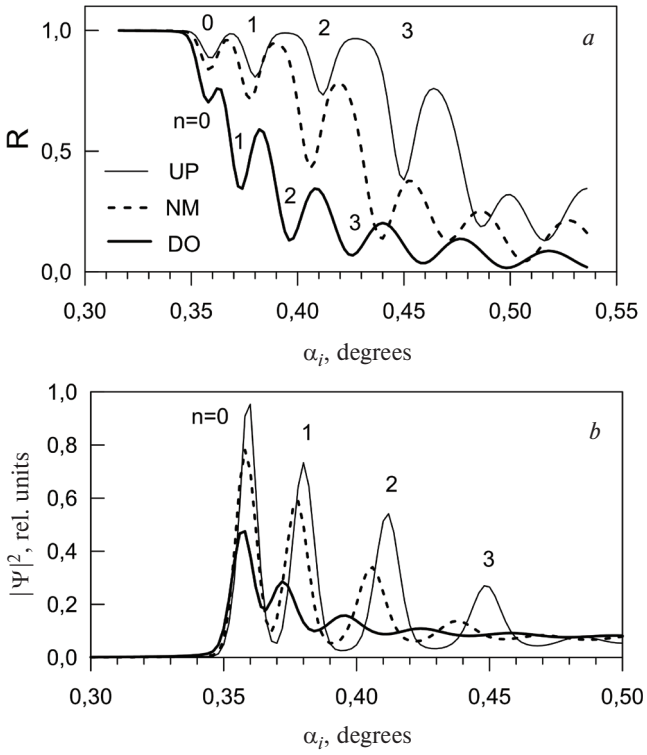

Fig. 5. Calculations: (a) Neutron specular reflection coefficient for UP polarization (thin line), DO polarization (thick line), and unpolarized NM beam (dashed line) as a function of the incident beam's grazing angle. (b) Square modulus of the neutron wavefunction for UP and DO polarizations and for the unpolarized NM beam as a function of the incident beam's grazing angle.

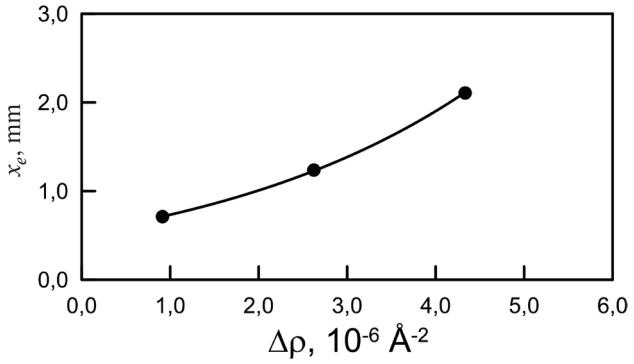

Fig. 6. Calculated neutron channeling length as a function of the waveguide potential well depth for different incident beam polarizations.

well depth, calculated for a neutron wavelength of 4.26 Å using Equation (5). The dots represent the calculated data, while the solid line represents an exponential fit. It is evident that the neutron channeling length increases exponentially with increasing waveguide potential well depth. Thus, preliminary calculations predict an exponential growth of both the square modulus of the neutron wavefunction and the neutron channeling length as the waveguide potential well depth increases.

3. EXPERIMENT

The experiments were conducted on the NREX polarized neutron reflectometer (FRM II reactor, Garching, Germany) [41]. The sample plane was positioned horizontally, allowing the boral block absorber to rest freely on the waveguide surface. The dimensions of the Py(20 nm)/Cu(140 nm)/Py(50 nm)//glass sample substrate were $30 \times 30 \times 5\ \text{mm}^3$, while the absorber block dimensions were $1 \times 1 \times 38\ \text{mm}^3$. The neutron wavelength was 4.26 Å. In polarized beam mode, the wavelength resolution was 1.5%, and the incident beam divergence was 0.006°. The angular resolution of the

^{3}He two-dimensional position-sensitive detector was 0.072°.

The polarization efficiency of the single supermirror polarizer was 97%, and it was used in the transmission geometry. An external magnetic field of 1.0 kOe was applied parallel to the sample plane to magnetize the permalloy films to saturation. The first aperture width was 0.25 mm, with a distance of 2200 mm from the first aperture to the sample and 2400 mm from the sample to the detector. The detector's spatial resolution was 3.0 mm. A second aperture, 0.7 mm wide, was placed 200 mm before the sample to reduce background noise.

The demagnetized state of the sample was achieved by applying an external magnetic field of +3 Oe along the film plane. This field value was determined from the hysteresis loop measured using the degree of polarization of the specularly reflected beam. During the determination of the neutron channeling length in the demagnetized waveguide, the unpolarized beam mode was used, with the polarizer removed from the beam path. The first aperture width was 0.35 mm, the neutron wavelength resolution was 2.0%, and the incident beam divergence was 0.009°.

Fig. 7a presents the neutron specular reflection coefficients for "+" polarization (light points) and "–" polarization (dark points) as a function of the incident beam's grazing angle. The solid lines show the fit results with the following parameters: layer thicknesses (nm), nuclear SLD (Å^{-2}), and magnetization of the layers (kG).

$$\text{PyO}(2.3\ \text{nm}, 7.67\cdot10^{-6}\ \text{Å}^{-2})/$$

$$/\text{Py}(19.5\ \text{nm}, 8.83\cdot10^{-6}\ \text{Å}^{-2}, 7.0\ \text{kG})/$$

$$/\text{Cu}(132.0\ \text{nm}, 6.58\cdot10^{-6}\ \text{Å}^{-2})/$$

$$/\text{Py}(48.0\ \text{nm}, 8.56\cdot10^{-6}\ \text{Å}^{-2}, 7.2\ \text{kG})//$$

$$//\text{glass}\ (2.63\cdot10^{-6}\ \text{Å}^{-2}).$$

The fit results indicate that the magnetization of the upper permalloy layer is 7.0 kG, while the magnetization of the lower layer is 7.2 kG. Fig. 7b shows the specular reflection coefficient for the unpolarized neutron beam reflected from the demagnetized sample. The fit with zero magnetization of the permalloy layers accurately describes the experimental data.

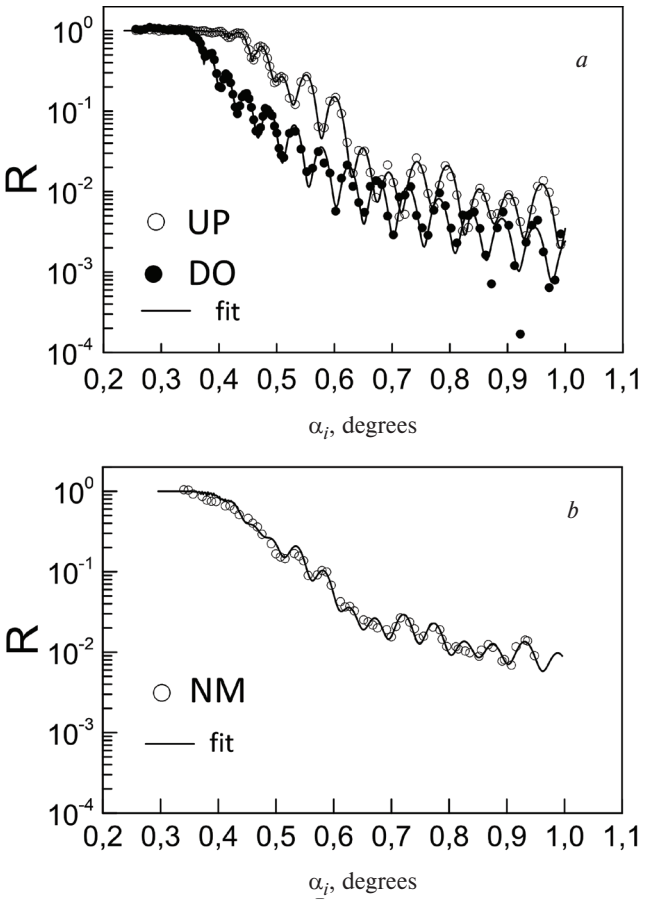

Fig. 7. Neutron specular reflection coefficients as a function of the incident beam's grazing angle (dots – experiment, lines – fit): (a) polarized beams UP and DO; (b) unpolarized beam NM.

Fig. 8a displays the neutron microbeam intensity without an absorber on the waveguide surface as a function of the incident beam's grazing angle for the initial polarizations "+" (light symbols) and "–" (dark symbols) with the fully illuminated waveguide surface. The resonance peaks are labeled with the corresponding resonance orders $n = 0, 1, 2\ldots$.

It can be seen that the microbeam intensity peak at the $n = 0$ resonance (background-subtracted) for the initial "–" polarization is approximately twice lower than the peak intensity for the "+" polarization. Higher-order resonances ($n = 1, 2, 3$) are clearly visible for the "+" polarization. For the "–" polarization, only a small peak at the $n = 1$ resonance is observed, significantly shifted to lower incident angles compared to the $n = 1$ resonance for the "+" polarization. The intensity of higher-order resonances for the "–" polarization is low, making the corresponding peaks barely visible.

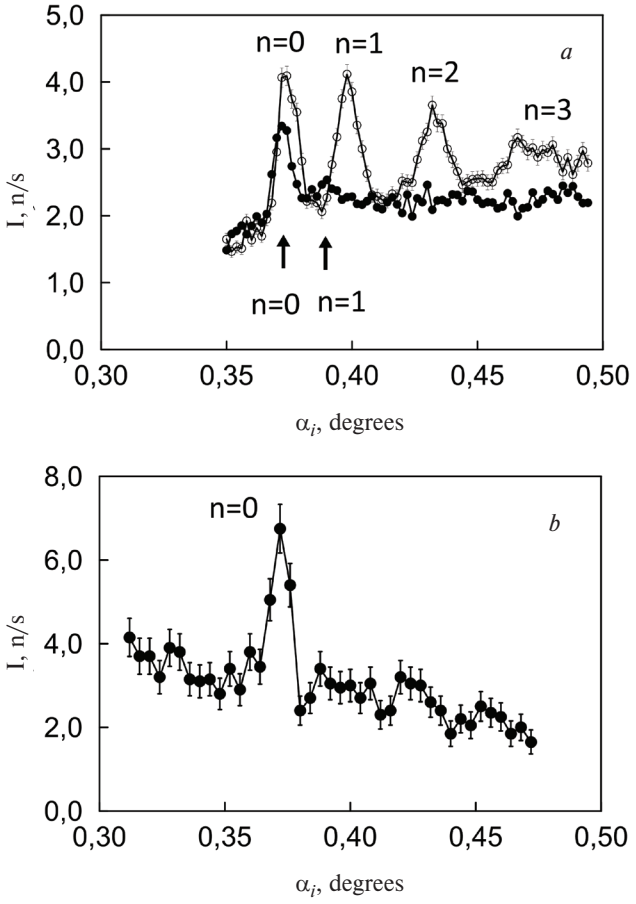

Fig. 8. Neutron microbeam intensity as a function of the incident beam's grazing angle: (a) incident beam polarization UP (light symbols) and DO (dark symbols); (b) unpolarized beam.

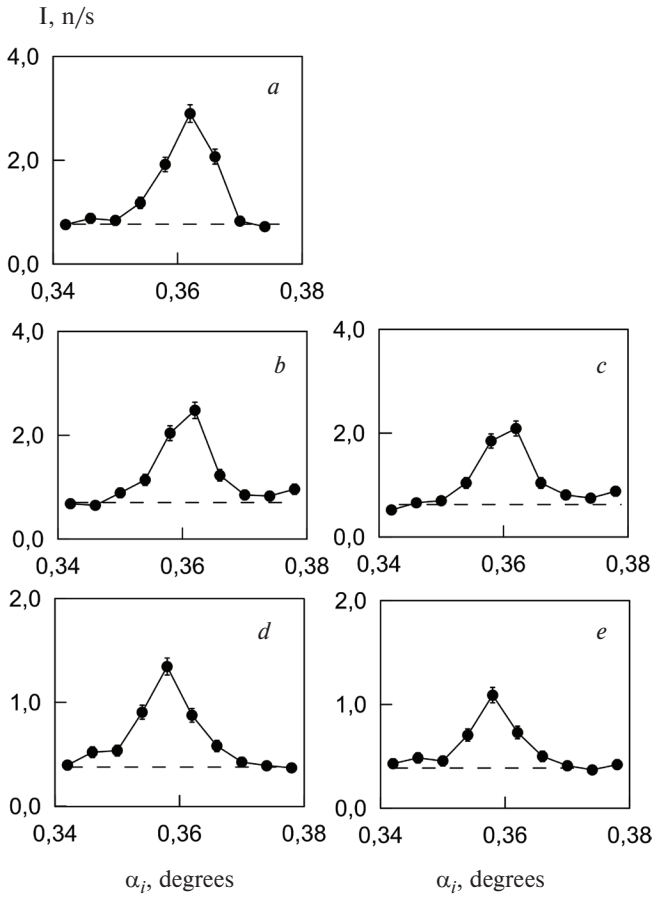

Fig. 9. Microbeam intensity for UP polarization as a function of the incident beam's grazing angle at different distances L between the front edge of the absorber on the surface and the output edge of the waveguide: (a) 1.0 mm; (b) 1.5 mm; (c) 2.5 mm; (d) 3.5 mm; (e) 4.0 mm. The dashed line indicates the background level. Data obtained from [38].

Fig. 8b shows the neutron microbeam intensity without an absorber on the surface of the demagnetized sample as a function of the incident angle of the unpolarized neutron beam. The peak corresponding to the $n = 0$ resonance is clearly visible. For normalization, the microbeam intensity $I(x = 0)$ is measured with the absorber placed at the very edge of the waveguide exit, when the waveguide surface is fully illuminated by the incident neutron beam. In this case, the main part of the specularly reflected beam is blocked by the absorber, reducing the background level near the microbeam by approximately 50%. This position corresponds to an absorber offset of $L = 1.0\,\text{mm}$ relative to the waveguide exit edge.

Fig. 9 presents the neutron microbeam intensity as a function of the grazing angle of the incident UP-polarized beam for different absorber positions relative to the waveguide exit edge: 1.0, 1.5, 2.5, 3.5, and 4.0 mm. These data were obtained and published in our previous work [38]. As the absorber moves away from the waveguide edge, the neutron microbeam intensity decreases systematically.Further studies were conducted on the demagnetized sample. Fig. 10 shows the intensity of the neutron microbeam as a function of the grazing angle of the incident unpolarized neutron beam (NM) for the absorber block positioned relative to the waveguide's output edge at 1.0, 2.3, and 2.7 mm. It can be observed that the microbeam intensity decreases as the distance from the waveguide's output edge to the absorber's front edge increases.

Fig. 11 presents the neutron microbeam intensity as a function of the grazing angle of the incident polarized DO beam for different absorber positions relative to the waveguide's output edge: 1.0, 1.7, 1.9, 2.2, and 2.4 mm. As seen, the microbeam

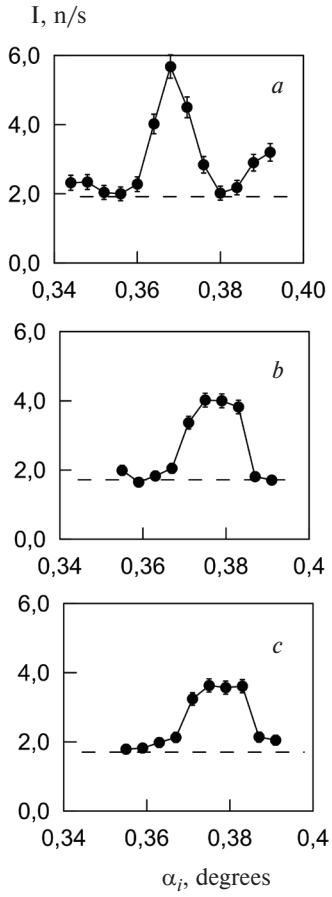

Fig. 10. Microbeam intensity in the unpolarized mode as a function of the incident beam's grazing angle at different distances L between the front edge of the absorber on the surface and the output edge of the waveguide: (a) 1.0 mm; (b) 2.3 mm; (c) 2.7 mm. The dashed line indicates the background level.

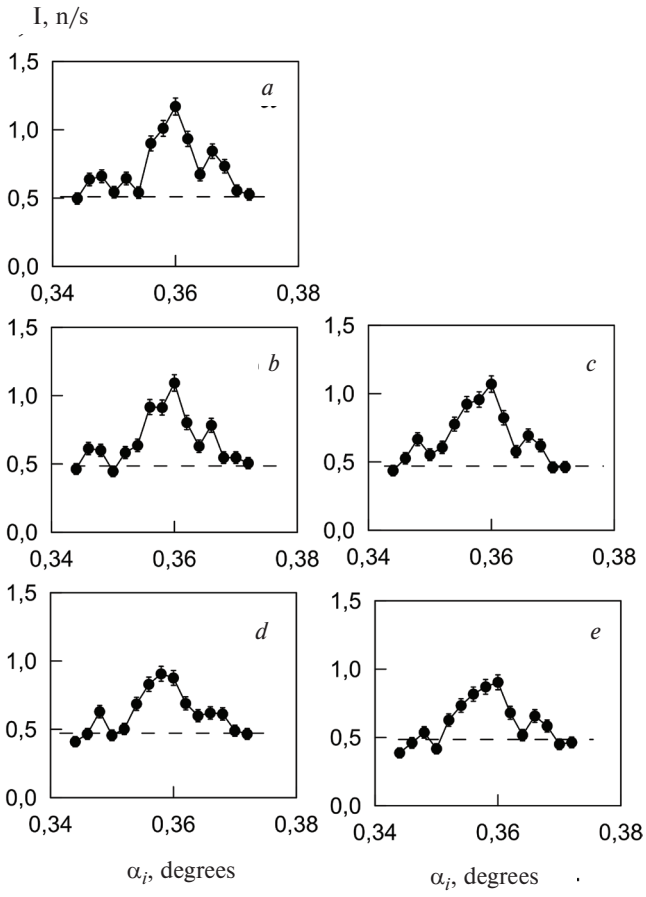

Fig. 11. Microbeam intensity for DO polarization as a function of the incident beam's grazing angle at different distances L between the front edge of the absorber on the surface and the output edge of the waveguide: (a) 1.0 mm; (b) 1.7 mm; (c) 1.9 mm; (d) 2.2 mm; (e) 2.4 mm. The dashed line indicates the background level.

intensity decreases with increasing distance from the waveguide output to the front edge of the absorber block.

Fig. 12 displays the normalized neutron microbeam intensity $I(L)/I(x=0)$ on a natural logarithmic scale as a function of the distance L from the waveguide's output edge to the absorber block's front edge (upper scale and light symbols) for the incident polarized UP beam (a), the unpolarized NM beam and demagnetized sample (b), and the polarized DO beam (c). The condition $I(L=1\,\text{mm}) = I(x=0)$ is taken into account.

It can be observed that the experimental points for $L > 1$ mm align along a straight line intersecting the 1.00 level at L. Vertical error bars represent statistical uncertainties in the neutron microbeam intensity. The data processing is performed as follows: the normalized intensity point at $L = 1$ mm is placed at the origin ($x = 0$) on the horizontal axis. Subsequently, all remaining points along the L coordinate (light symbols) are shifted by a single value L along the horizontal axis so that the line through all $x = L - L'$ coordinate points (dark symbols and lower scale) passes through the origin $x = 0$. The shift value L' depends on the accuracy of the initial absorber block positioning relative to the waveguide's output edge and the size of the air gap between the absorber and the waveguide surface. Consequently, the line $\ln[I(x)/I(x=0)] = -x/x_e$ intersects the 0.37 level on the vertical axis at the point corresponding to the experimental neutron channeling length. The uncertainty in the neutron channeling length is determined by the extreme trajectories passing through the experimental points, considering the statistical error of the microbeam intensity. It is noteworthy that the longest neutron channeling length is observed for the UP-polarized incident

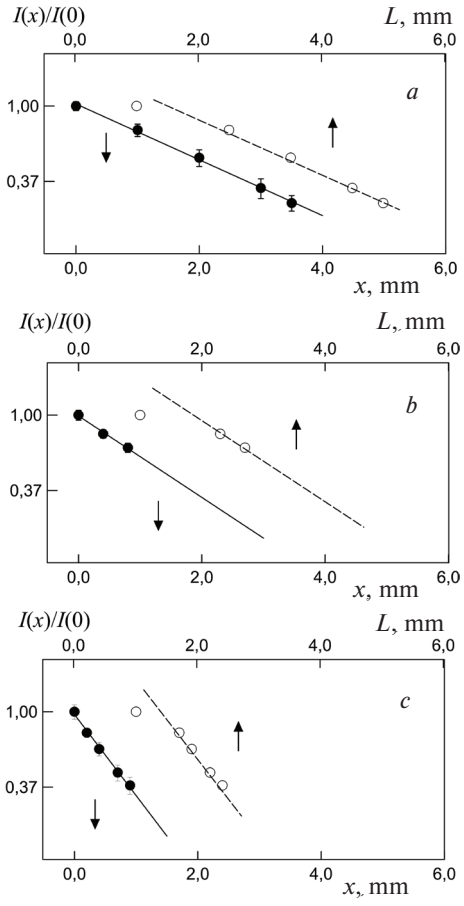

Fig. 12. Normalized microbeam intensity on a natural logarithmic scale as a function of the distance L from the output edge of the waveguide to the front edge of the absorber block (upper scale and light symbols) and the length of the non-illuminated surface area of the waveguide x (lower scale and dark symbols) for different polarization values: (a) UP, (b) NM, (c) DO.

beam (Fig. 12a). The channeling length decreases for the unpolarized NM beam and the demagnetized sample (Fig. 12b). The shortest channeling length is observed for the DO polarization (Fig. 12c).

The illuminated surface area under the absorber Δx in Fig. 3 corresponds to the point L' when the absorber block just begins to partially cover the waveguide surface from the incident beam. From the equality $\Delta x = L'$, the air gap under the block can be estimated as $h \approx \alpha_i \Delta x$.

Fig. 13 shows the experimental neutron channeling length as a function of the waveguide's scattering length density (SLD) depth $\Delta\rho = \rho_1 - \rho_2$. The points represent the experimental data, while the line represents the least-squares exponential fit. It is evident that the experimental data follow an exponential dependence, qualitatively confirming

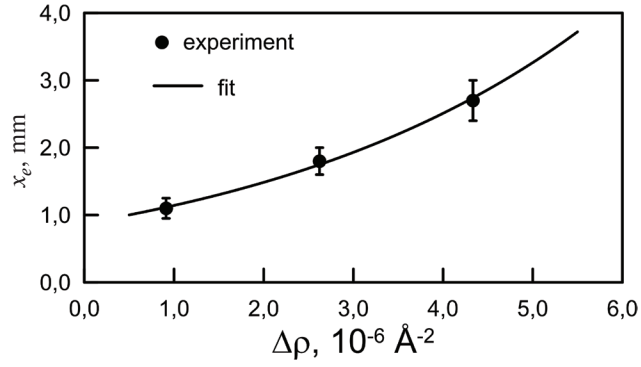

Fig. 13. Neutron channeling length as a function of the waveguide's SLD depth for different incident beam polarizations. Points represent experimental data, and the line represents the least squares exponential fit.

the preliminary channeling theory calculations. The quantitative comparison between theory and experiment depends on the precise determination of the actual structural parameters (oxide layer thickness, layer thicknesses, SLD, and layer magnetization). However, minor deviations of the experimentally obtained structure parameters from nominal values should not affect the overall trend of the neutron channeling length's dependence on the waveguide potential well depth. The channeling theory was previously validated experimentally [35]. The calculated neutron channeling length, with refined Fe/Cu/Fe//glass waveguide parameters, matched the experimentally obtained value within statistical error limits.

4. DISCUSSION OF RESULTS

The examined magnetic waveguide Py/Cu/Py can be utilized in two directions. First, as a polarizer for generating a polarized neutron microbeam in experiments studying magnetic microstructures. From the intensity ratio of the UP and DO microbeam polarizations for the resonance of order $n = 0$ (see Fig. 8a), it follows that the waveguide's polarization efficiency is 0.3. The review [37] provides a detailed discussion of various polarizing and non-polarizing magnetic waveguides. For example, the Fe(20 nm)/Cu(140 nm)/Fe(50 nm)// glass waveguide has a polarization efficiency of 0.6 for the $n = 0$ resonance. The polarization efficiency of the magnetic waveguide Fe(20 nm)/Co(150 nm)/ Fe(50 nm)//Si reaches 1.0.

Magnetic waveguides have a significant drawback. Due to the high divergence of the microbeam, the

sample under study must be located approximately 1 mm from the waveguide's exit. In such an experimental setup, it is challenging to separate the magnetic field on the waveguide from the field on the sample. The most practical solution is a combination of a polarized neutron reflectometer and a non-magnetic waveguide [42]. In this configuration, high polarization of the microbeam is achieved conventionally, and the magnetic field on the sample does not affect the operation of the non-magnetic waveguide. This non-magnetic waveguide setup was used in the experiment with the magnetic microwire [2, 8]. Thus, non-magnetic waveguides have an advantage over magnetic ones in experiments studying magnetic microstructures using polarized neutron microbeams.

The second application of magnetic neutron waveguides is for controlling the chain reaction of uranium fission. The idea proposed in [31] suggests that by remagnetizing the external magnetic layers with an applied magnetic field, the neutron density in the middle non-magnetic layer can be altered. If uranium is placed inside the non-magnetic layer, the uranium fission reaction can be controlled using an external magnetic field. Suitable candidates for this method include the Py/Cu/Py and Fe/Cu/Fe magnetic waveguides. The preferred choice is the Fe/Cu/Fe waveguide, which has twice the polarization efficiency. However, in this study, we investigated the Py/Cu/Py waveguide. For the DO polarization of the incident beam, the waveguide retains a shallow potential well of the scattering length density (SLD), enabling experimental measurement of the neutron channelling length for this polarization. Clearly, non-magnetic waveguides are not suitable for controlling the chain reaction, as they do not respond to the magnetic field.

5. CONCLUSION

This study investigated the Py/Cu/Py//glass neutron waveguide with external magnetic layers. The magnitude of the SLD of the magnetic layer depends on the sign of the incident neutron beam polarization. Preliminary calculations based on the theory of resonances in layered nanostructures showed that the square of the neutron wave function modulus inside the waveguide increases with the depth of the SLD potential well. Calculations using the theory of neutron channelling in planar waveguides predicted an exponential growth of the neutron channelling length with increasing SLD potential well depth.

The neutron channelling length was experimentally determined for the UP and DO polarization of the incident neutron beam for a sample magnetized to saturation and for the unpolarized incident beam for a fully demagnetized sample (non-magnetic or NM mode). The results showed that the neutron channelling length increases exponentially with the depth of the SLD potential well. Thus, the experimental results confirm the predictions of the neutron channelling theory in layered nanostructures.

ACKNOWLEDGMENTS

The authors thank T. Keller and F. Radu for useful discussions.

REFERENCES

1. *F. Ott*, *Focusing Optics for Neutrons*. In: Modern Developments in X-Ray and Neutron Optics. Springer Series in Optical Sciences. 2008. Vol. **137**. pp. 113–134.
2. *F. Ott, S. Kozhevnikov, A. Thiaville, J. Torrejon, M. Vazquez*, Nucl. Instrum. and Meth. A **788**, 29 (2015).
3. *F. Pfeiffer, V. Leiner, P. Høghøj, I. Anderson*, Phys. Rev. Lett. **88**, 055507 (2002).
4. *S. V. Kozhevnikov, A. Rühm, F. Ott, N. K. Pleshanov, J. Major*, Physica B **406**, 2463 (2011).
5. *S. V. Kozhevnikov, V. K. Ignatovich, Yu. V. Nikitenko, F. Ott, A. V. Petrenko*, JETP Letters **102**, 1 (2015).
6. *S. V. Kozhevnikov, V. D. Zhaketov, F. Radu*, JETP **127**, 593 (2018).
7. *S. V. Kozhevnikov, V. D. Zhaketov, T. Keller, Yu. N. Khaydukov, F. Ott, F. Radu*, Nucl. Instrum. and Meth. A **915**, 54 (2019).
8. *S. V. Kozhevnikov, F. Ott, J. Torrejón, M. Vázquez, A. Thiaville*, Phys. Solid State **56**, 57 (2014).
9. *F. Radu, V. K. Ignatovich*, Physica B **292**, 160 (2000).
10. *S. V. Kozhevnikov, V. K. Ignatovich, A. V. Petrenko, F. Radu*, JETP **123**, 950 (2016).
11. *Yu. V. Nikitenko*, Phys. Part. Nucl. **40**, 890 (2009).
12. *H. Zhang, P. D. Gallagher, S. K. Satija, R. M. Lindstrom, R. L. Paul, T. P. Russell, P. Lambooy, E. J. Kramer*, Phys. Rev. Lett. **72**, 3044 (1994).
13. *V. L. Aksenov, Yu. V. Nikitenko, F. Radu, Yu. M. Gledenov, P. V. Sedyshev*, Physica B **276–278**, 946 (2000).
14. *V. D. Zhaketov, K. Hramco, A. V. Petrenko, Yu. N. Khaydukov, A. Csik, Yu. N. Kopatch, N. A. Gundorin, Yu. V. Nikitenko, V. L. Aksenov*, Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques **15**, 549 (2021).
15. *V. D. Zhaketov, A. V. Petrenko, S. N. Vdovichev, V. V. Travkin, A. Csik, Yu. N. Kopatch, Yu. M. Gledenov,*

E. Sansarbayar, N. A. Gundorin, Yu. V. Nikitenko, V. L. Aksenov, Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques **13**, 478 (2019).

16. *Yu. V. Nikitenko, A. V. Petrenko, N. A. Gundorin, Yu. M. Gledenov, V. L. Aksenov*, Crystallogr. Rep. **60**, 466 (2015).

17. *V. L. Aksenov, V. D. Zhaketov, Yu. V. Nikitenko*. Phys. Part. Nucl. **54**, 756 (2023).

18. *V. L. Aksenov, Yu. V. Nikitenko, S. V. Kozhevnikov, F. Radu, R. Kruijs, T. Rekveldt*, Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques **16**, 1225 (2001).

19. *Yu. Khaydukov, A. M. Petrzhik, I. V. Borisenko, A. Kalabukhov, D. Winkler, T. Keller, G. A. Ovsyannikov, B. Keimer*, Phys. Rev. B **96**, 165414 (2017).

20. *Yu. N. Khaydukov, D. Lenk, V. Zdravkov, R. Morari, T. Keller, A. S. Sidorenko, L. R. Tagirov, R. Tidecks, S. Horn, B. Keimer*, Phys. Rev. B **104**, 174445 (2021).

21. *M. Wolff, A. Devishvili, J. A. Dura, F. A. Adlmann, B. Kitchen, G. K. Pálsson, H. Palonen, B. B. Maranville, Ch. F. Majkrzak, B. P. Toperverg*, Phys. Rev. Lett. **123**, 016101 (2019).

22. *S. V. Kozhevnikov, F. Ott, E. Kentzinger, A. Paul*, Physica B **397**, 68 (2007).

23. *S. V. Kozhevnikov, F. Ott, A. Paul, L. Rosta*, Eur. Phys. J. Special Topics **167**, 87 (2009).

24. *E. Kentzinger, U. Rucker, B. Toperverg, T. Bruckel*, Physica B **335**, 89 (2003).

25. *F. Radu, A. Vorobiev, J. Major, H. Humblot, K. Westerholt, H. Zabel*, Physica B **335**, 63 (2003).

26. *L. Guasco, Y. N. Khaydukov, S. Pütter, L. Silvi, M. Paulin, T. Keller, B. Keimer*, Nature Communications **13**, 1486 (2022).

27. *A. Perrichon, A. Devishvili, K. Komander, G. K. Pálsson, A. Vorobiev, R. Lavén, M. Karlsson, M. Wolff*, Phys. Rev. B **103**, 235423 (2021).

28. *S. P. Pogossian, H. Le Gall, A. Menelle*, J. Magn. Magn. Mater. **152**, 305 (1996).

29. *S. V. Kozhevnikov, Yu. N. Khaydukov, T. Keller, F. Ott, F. Radu*, JETP Letters **103**, 36 (2016).

30. *S. V. Kozhevnikov, V. D. Zhaketov, T. Keller, Yu. N. Khaydukov, F. Ott, Chen Luo, Kai Chen, F. Radu*, Nucl. Instrum. and Meth. A **927**, 87 (2019).

31. *S. P. Pogossian*, J. Appl. Phys. **102**, 104501 (2007).

32. *V. K. Ignatovich, F. Radu*, Phys. Rev. B **64**, 205408 (2001).

33. *S. V. Kozhevnikov, T. Keller, Yu. N. Khaydukov, F. Ott, F. Radu*, Nucl. Instrum. and Meth. A **875**, 177 (2017).

34. *V. L. Aksenov, Yu. V. Nikitenko*, Physica B **297**, 101 (2001).

35. *S. V. Kozhevnikov, V. K. Ignatovich, F. Ott, A. Rühm, J. Major*, JETP **117**, 636 (2013).

36. *Yu. V. Nikitenko, V. V. Proglyado, V. L. Aksenov*, Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques **8**, 961 (2014).

37. S. V. Kozhevnikov. Phys. Particl. Nucl. **50**, 300 (2019).

38. *S. V. Kozhevnikov, V. D. Zhaketov, Yu. N. Khaydukov, F. Ott, F. Radu*, JETP **125**, 1015 (2017).

39. *S. V. Kozhevnikov, T. Keller, Yu. N. Khaydukov, F. Ott, F. Radu*, JETP **128**, 504 (2019).

40. *S. V. Kozhevnikov, Yu. N. Khaydukov, F. Ott, F. Radu*, JETP **126**, 592 (2018).

41. *Yu. N. Khaydukov, O. Soltwedel, T. Keller*, J. Large Scale Research Facilities A **38**, 1–4 (2015).

42. *S. V. Kozhevnikov, A. Rühm, J. Major*, Crystallography Reports **56**, 1207 (2011).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 76–82

SOLIDS AND LIQUIDS

DIAELASTIC EFFECT IN ALUMINUM DUE TO LASER ULTRAVIOLET NANOSECOND LASER PULSE IRRADIATION

© 2025 G. V. Afonin[a], V. Yu. Zheleznov[b], T. V. Malinskiy[b], S. I. Mikolutzkiy[b], V. E. Rogalin[b], Yu. V. Khomich[b], N. P. Kobelev[c], V. A. Khonik[a,*]

[a]Voronezh State Pedagogical University (VSPU), Voronezh, Russia
[b]Institute for Electrophysics and Electric Power RAS (IEE RAS), Saint Petersburg, Russia
[c]Osipyan Institute of Solid State Physics RAS (ISSP RAS), Chernogolovka, Russia
*e-mail: v.a.khonik@yandex.ru

Received July 03, 2024

Revised October 07, 2024

Accepted October 08, 2024

Abstract. It is shown that irradiation of single-crystal aluminum with nanosecond ultraviolet pulses, causing its surface melting, leads to a decrease in all resonance frequencies of the ultrasonic vibration spectrum of the sample. The shear modulus decreases from 0.87% to 1.45% with an increase in the incident radiation density from 1.1 J/cm^2 to 5.3 J/cm^2. Subsequent heating to pre-melting temperatures causes the shear modulus to be restored to its original value. The hypothesis is argued that the discovered diaelastic effect is due to interstitial atoms in a dumbbell configuration, formed in the surface layer as a result of melting and preserved in this layer in a solid state due to the high rate of its cooling. The possibilities of another interpretation are discussed.

Keywords: *Nanosecond laser pulses, surface melting, ultrasound spectroscopy, shear modulus, interstitial atoms, dumbbell configuration*

DOI: 10.31857/S00444510250107e7

1. INTRODUCTION

Laser processing is one of the most promising and in-demand methods for modifying the physical properties of materials. The use of short laser pulses allows achieving high heating and cooling rates of the near-surface layer of the material [1]. It turns out that the behavior of solids during rapid processes changes significantly. This can lead to a fundamental alteration of properties, enabling the creation of materials with new mechanical, electrical, and optical characteristics. Pulsed lasers serve as a convenient tool for experimental research in the development of new materials and the study of their properties [2]. The results obtained using them provide broad opportunities for an in-depth understanding of such phenomena as phase transitions, recrystallization, formation of structural defects, amorphization, etc. [4, 3].

A particular interest lies in the use of laser pulses with a duration of about 10 ns and an energy density of several J/cm^2, leading to the melting of the surface layer of a substance within the duration of the pulse. The cooling of this layer occurs over a time comparable to the pulse duration [5, 6], and the cooling rate when using nanosecond pulses can reach 10^8 K/s [7]. In the case of a pure metal, such a cooling rate is insufficient to form an amorphous layer (for example, pure vanadium and tantalum vitrify at a quenching rate of approximately 10^{12} K/s [8]; there is reason to believe that vitrification of monatomic metals is feasible in principle [9]). However, it is evident that its defect structure will undergo significant changes. Consequently, macroscopic elastic characteristics may also change, as the mechanical properties of crystals are largely determined by their defect structure. This experimental scheme was implemented in the present work. The object of study was chosen to be pure monocrystalline aluminum, subjected to nanosecond pulses of an ultraviolet laser, which induced melting of the surface layer and its subsequent rapid cooling (quenching).

The initial motivation for this work was as follows. Granato's well-known interstitialcy theory argues that metal melting results from the

avalanche-like generation of interstitial atoms in a dumbbell configuration (interstitial dumbbells), leading to a significant reduction in shear modulus and destabilization of the crystal lattice [10, 11]. The application of this theory to the case of multicomponent metallic glasses yields very good results, allowing a quantitative interpretation of changes in their properties during heat treatment in a solid amorphous state and tracing the connection between these changes and the properties of the melt and the parent crystal [11]. However, information on the applicability of these concepts to pure metals remains quite limited.

Firstly, it has been shown that monocrystalline aluminum in the premelting region exhibits a measurable diaelastic effect – a reduction in the shear modulus beyond the standard purely anharmonic decrease, indicating a significant increase in the concentration of interstitial dumbbells as the melting temperature is approached [12]. A similar situation is observed in polycrystalline indium [13]. Secondly, it was established that the observed premelting nonlinear increase in the heat capacity of aluminum can also be attributed to the intensive generation of interstitial dumbbells [14]. Finally, thirdly, about 70% of the total melting entropy of aluminum (and, accordingly, the heat of fusion) observed in experiments can be interpreted as the result of interstitial dumbbell generation at the melting temperature [15].

Based on this information, it was hypothesized that laser surface melting of aluminum would cause a significant increase in the concentration of interstitial dumbbell-type defects in the melt, and subsequent rapid cooling would "freeze" them in the solid crystalline state. The frozen interstitial dumbbells would induce a measurable diaelastic effect, the magnitude of which could indicate the concentration of these defects in the melt. However, other mechanisms of the diaelastic effect are also possible, as discussed below.

2. EXPERIMENTAL METHODOLOGY

Monocrystals of aluminum with a purity of 99.996%, grown using a modified Bridgman method with orientation 100 along the growth axis, were studied. Orientation control was performed using the X-ray method [12]. Samples in the shape of a cube with a side length of 3 mm were then prepared from the grown crystal using electrical discharge machining. Each face of the cube was perpendicular to the [100] direction. The samples were then processed on a grinding machine with 1200-grit abrasive and annealed by heating to 923 K followed by slow cooling.

Sample processing was performed using a scanning laser beam from an Optolette HR2731 (OPOTEC Inc.), which generated radiation pulses with a wavelength of 355 nm, a duration of approximately 10 ns, an output energy of up to 2 mJ, and a pulse repetition rate of 100 Hz. The laser was calibrated using a Nova II energy meter (Ophir Optronics Solutions Ltd.) with a pyroelectric detector PE50-SH-V2. The laser spot size in the sample surface plane was determined using a standard method [16] by measuring the area of imprints left by laser pulses on a reference aluminum plate. The characteristic diameter of the laser spot in these experiments was 180 μm. Surface processing of the sample was performed using a two-coordinate table, enabling sample movement along a "*serpentine*" trajectory at a speed of 3 mm/s with a line spacing of *25* μ*m*, ensuring that adjacent laser spots overlapped with a coverage coefficient of no less than 98%. Each area of the surface was exposed to 30 laser pulses. The energy densities on the surface of the processed samples exceeded the ablation threshold and were 1.1 J/cm^2, 2.4 J/cm^2, and 5.3 J/cm^2. The exceeding of the ablation threshold was visually observed as an accompanying plasma plume and confirmed through electron microscopy images of the irradiated sample surfaces. All six faces of each cubic sample were sequentially processed. The surfaces of the samples before and after laser exposure were studied using a multi-beam optical profilometer Zygo NewView 7300.

The irradiated samples were then examined using resonant ultrasound spectroscopy (RUS) on a setup similar to that described in [17]. The excitation and detection of ultrasonic vibrations were carried out using piezoelectric transducers, which pressed the opposite vertices of the cubic sample. A special lever-type system minimized the axial pressure of the piezoelectric transducers on the sample, ensuring that the measured resonance spectrum was close to the natural one (i.e., determined only by the sample properties and its geometry). An advanced hardware-software RUS signal processing firmware enabled the registration of sample resonance frequencies with high precision, down to ppm levels. A total of

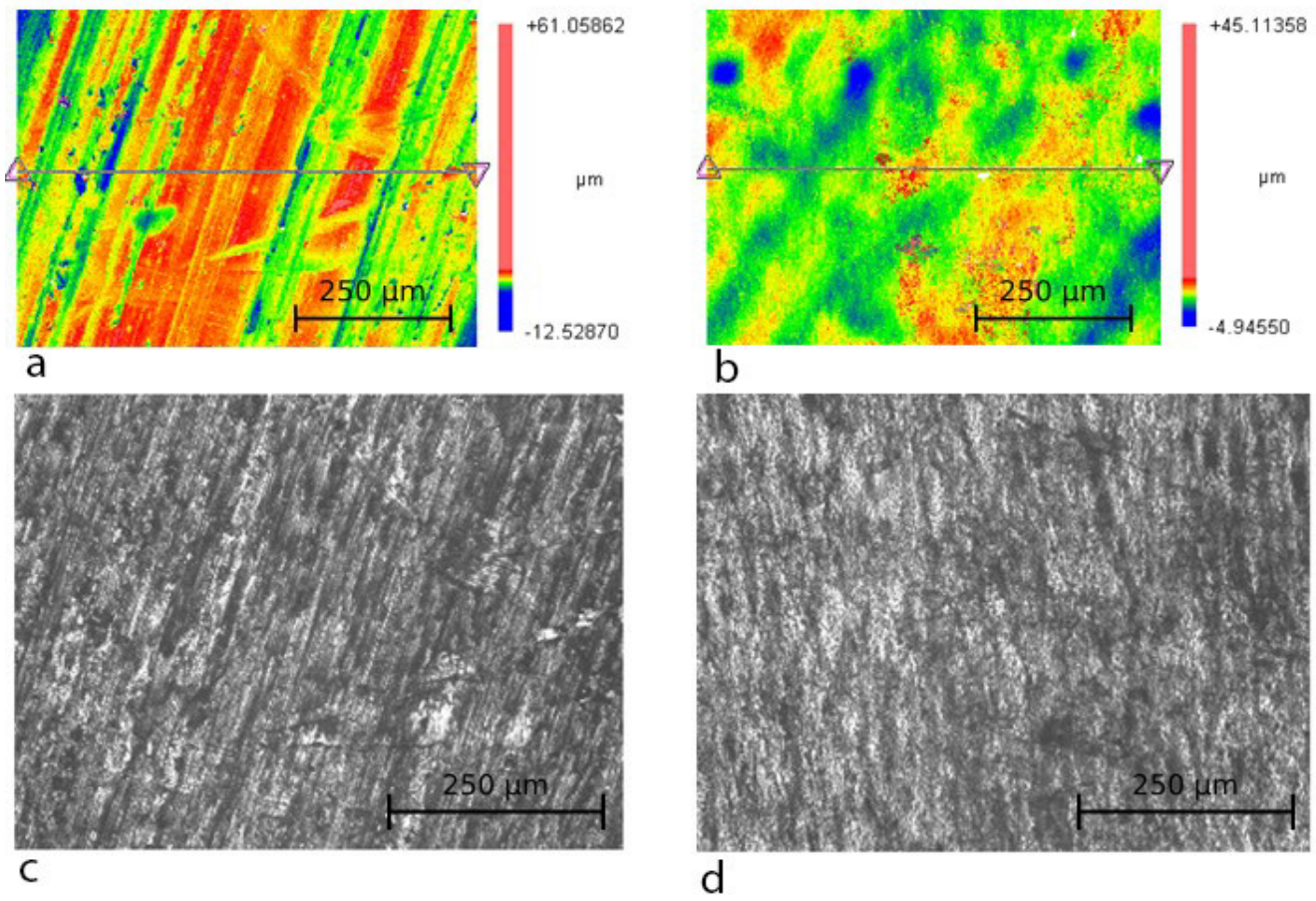

Fig. 1. (Color online) 2D profilograms (a, b) and surface micrographs (c, d) (optical profilometer Zygo NewView 7300) of sample S5 in the initial state (a, c) and after treatment with 30 UV laser pulses with an energy density of 5.3 J/cm² (b, d).

five samples were studied in their initial state and after various laser treatments. The research results are illustrated below with data for three of them. Additionally, it should be noted that the time from sample irradiation to RUS measurement was several weeks.

3. RESULTS

Fig. 1 shows, as an example, 2D profilograms (a, b) and surface micrographs (c, d) of sample S5 in its initial state (a, c) and after laser exposure with 30 pulses at an energy density of $W_p = 5.3$ J/cm² (b, d). Linear morphological features in the initial state (a, c) correspond to abrasive processing traces. After laser exposure, these features disappear (b, d), and irregular roughness is observed on the surface, with a height comparable to that before irradiation. Detailed studies using a scanning electron microscope (SEM) revealed clear evidence of surface melting. A similar situation was observed for treatments with other laser energy densities.

The RUS spectra of the studied samples over the full range of resonance frequencies (500kHz<f<1300 kHz) contain 10–12 peaks corresponding to different elastic moduli and various interference modes due to the non-parallelism of the sample faces and other geometric defects. Fig. 2 shows the initial sections of the RUS spectra of samples S1, S3, and S5 after laser exposure at 1.1 J/cm² (*a*), 2.4 J/cm² (*b*), and 5.3 J/cm² (*c*), followed by annealing via heating to 850 K and slow cooling, demonstrating the presence of several resonances. The differences in the absolute values of the resonance frequencies for different samples are due to variations in their geometric dimensions. It is evident that the resonance frequencies of the irradiated samples in all cases are slightly lower than those observed after annealing, while the resonance peak heights

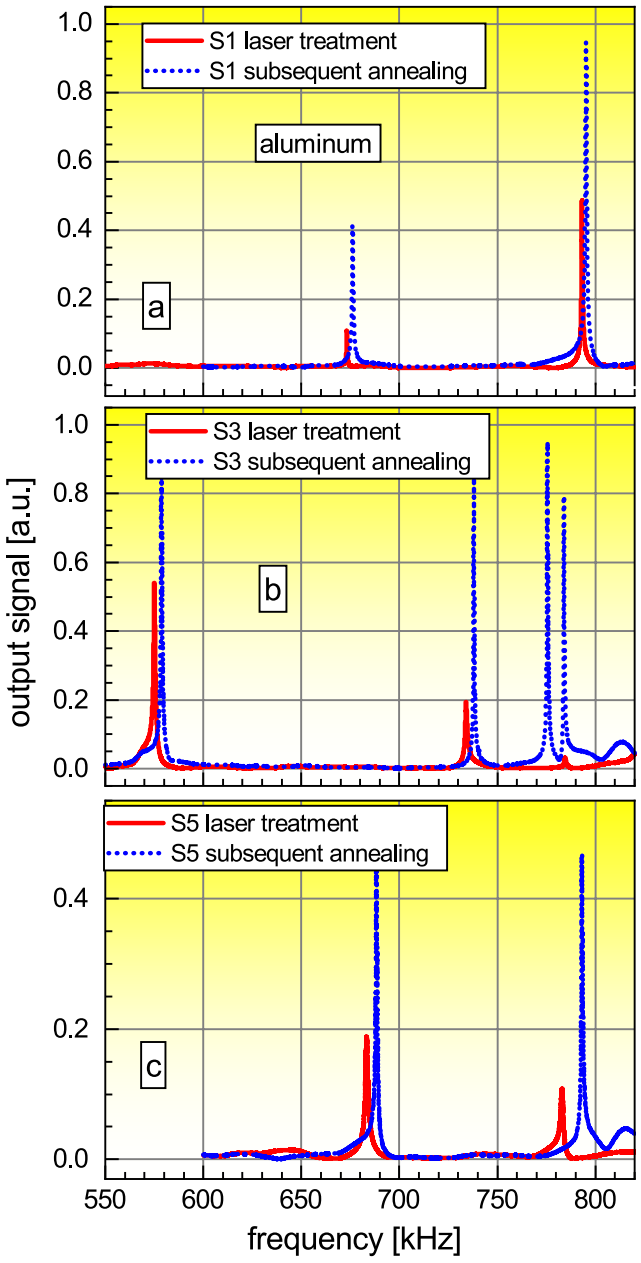

Fig. 2. (Color online) Initial sections of the resonant ultrasound spectroscopy spectra of aluminum samples S1 (a), S3 (b), and S5 (c), treated with UV laser pulses at energy densities of 1.1, 2.4, and 5.3 J/cm², respectively. The spectra of the same samples after heating to 850 K at a rate of 3 K/min are also shown. It is evident that the resonance frequencies increase in all cases after annealing.

significantly increase after annealing, which unequivocally indicates a reduction in sample defectiveness. As is known, the lowest resonance frequency corresponds to pure shear vibrations and is controlled by the shear modulus (C_{44}) [17]. This modulus (denoted as G hereafter) is of primary interest in this study.

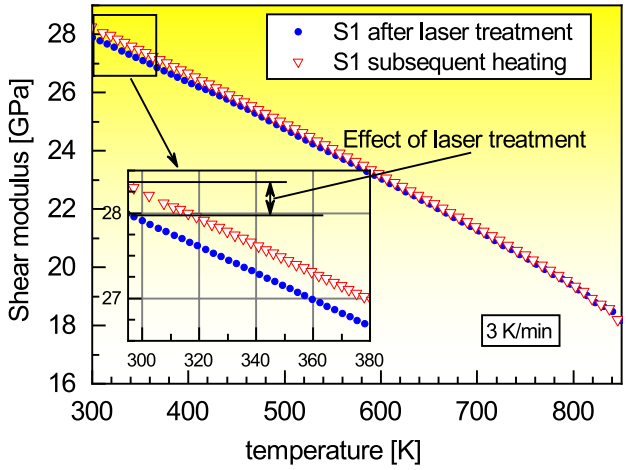

Fig. 3. (Color online) Temperature dependence of the shear modulus of sample S1 after laser exposure and subsequent reheating to 850 K. The inset shows the initial sections of these dependencies. Arrows indicate the effect of laser exposure on the shear modulus of the annealed sample. It is evident that heating to 850 K eliminates the influence of laser exposure on the shear modulus.

Table. Resonance frequencies of shear vibrations after laser exposure (f_{irr}) and subsequent heating to 850 K (f_{ann}), as well as the corresponding relative changes in the shear modulus ($\Delta G/Go$) for samples S1, S3, and S5, subjected to laser exposure at the specified energy densities (W_p). The error in determining the resonance frequencies and their changes after annealing is approximately 5 ppm.

Sample No.	W_p, J/cm^{-2}	F_{irr}, kHz	F_{ann}, kHz	$\Delta G/G$
S_1	1.1	673.20	676.16	–0.0087
S_3	2.4	573.98	578.17	–0.0127
S_5	5.3	683.23	688.24	–0.0145

The table presents the resonance frequencies corresponding to the shear modulus for samples S1, S3, and S5 after laser exposure (f_{irr}) and subsequent heating to 850 K (f_{ann}) for three energy densities (W_p). The relative changes in the shear modulus, calculated as

$$\Delta G / G = f_{irr}^2 / f_{ann}^2 - 1.$$

are also shown.

As seen, the shear modulus of sample S1 after irradiation at W_p = 1.1 J/cm² is by 0.87% lower than after annealing. The decrease in the modulus after laser exposure increases with the energy of the incident radiation, reaching 1.45% for sample S5 at W_p = 5.3 J/cm². This is the diaelastic effect discussed in this work. Notably, this effect occurs not only for the shear modulus but also for all other elastic moduli, as all resonance frequencies in the RUS

spectrum decrease as a result of laser exposure (see Fig. 2). It is important to emphasize that no similar data are known in the literature.

4. DISCUSSION

Metals, including the studied aluminum, absorb light through the transfer of photon energy to the electronic component of the skin layer, which has a thickness of approximately 10 nm [19]. The transfer of energy from the electronic subsystem to phonons takes several picoseconds. The heating time is approximately equal to the duration of the laser pulse and is of the order of 10 ns. At any of the applied laser energy densities W_p, surface melting of aluminum occurs, as confirmed by the aforementioned SEM observations of the irradiated sample surfaces. Simultaneously with the surface heating process, heat dissipation occurs due to thermodiffusion.

The characteristic heating depth during the laser pulse can be estimated as:

$$L = 2\sqrt{\alpha\tau}. \quad (1)$$

With $\alpha = 9.7 \times 10^{5}$ m²/s as the thermal diffusivity of aluminum and a laser pulse duration of 10 ns, using Equation (1), we obtain a heating depth of $L \approx$ 2μm. After the laser pulse ends, the cooling process of the heated region begins. The cooling time can be estimated from the equation [6]:

$$t_c = \frac{4L^2}{\alpha\pi^2}\ln\left(\frac{8T_m}{T_0\pi^2}\right), \quad (2)$$

where $T = 933$ K is the melting temperature of aluminum, and $T_0 = 300$ K is the initial temperature. Using Equation (2), the complete cooling time of the surface after laser exposure by a nanosecond pulse is $t_c \approx 25$ ns. The cooling (quenching) rate from the liquid state can be estimated as $T_v/t_c \approx 2 \cdot 10^{11}$ K/c, where $T_v = 2792$K is the boiling temperature of aluminum. As noted earlier, this cooling rate is insufficient for amorphization of the pure metal. Thus, within about 25 ns, the surface layer undergoes a phase transition from the crystalline state to the liquid state and back.

Following the concept of melting of simple metals outlined in the introduction, we assume that melting of the surface layer results in a high concentration of interstitial dumbbell defects, which, due to subsequent rapid quenching, become "frozen" into the crystalline structure.

The key feature of an interstitial dumbbell is that an externally applied alternating mechanical stress induces oscillatory motion of 20–30 atoms near its core (the atomic structure of this defect is shown in [11, 20]), leading to significant inelastic deformation and a corresponding reduction in the shear modulus [11, 21].

For the shear modulus (G) in the presence of interstitial dumbbells with a concentration c, the interstitialcy theory gives the relation [10, 21]:

$$G = G_0 \exp(-\alpha\beta_i c_i), \quad (3)$$

where G_0 is the shear modulus of the defect-free crystal, $\alpha \approx 1$ is a dimensionless constant, and β_i is the dimensionless shear susceptibility. Equation (3) shows that if the constant β_i is known, then by knowing the shear modulus of the defective crystal, one can estimate the concentration of interstitial dumbbells c_i, and vice versa.

A rough estimate using the Reuss approximation shows that the relative change in the shear compliance of the entire sample ($\Delta S/S$) is related to the compliance change of the molten layer ($\Delta S_{irr}/S_0$) by:

$$\Delta S/S = (\Delta S_{irr}/S)(\Delta V/V),$$

where $\Delta V/V$ is the ratio of the molten layer volume to the sample volume. As noted earlier, the sample melts to a characteristic depth of $L = 2$ μm. The fraction of the molten part of the cubic sample with an edge length a, having six faces, is

$$\Delta V / V = 6L / a = 6 \cdot 2 \cdot 10^{-6} / (2.2 \cdot 10^{-3}) \approx 6 \cdot 10^{-3}.$$

Since for small changes in shear elasticity, $\Delta S/S = -\Delta G/G$, from table data, we find that ΔS/S ranges from 0.009 to 0.014, depending on the laser energy density. Since shear compliance is the inverse of the shear modulus, the values of $\Delta S_{irr}/S$, range from 1.5 to 2.3, corresponding to S_{irr} values ranging from 2.4S to 3.4S. Using Equation (3), for the shear compliance of the irradiated crystal, we take

$$S_{irr} = \exp(\alpha\beta_i c_i) / G_0,$$

where for interstitial dumbbells in aluminum $\alpha\beta_i = 27$ [22]. Thus, the concentration of interstitial dumbbells c, which provides such compliance in the molten layer, should be 0.033 to 0.045.

Given the approximate nature of the initial data, this estimate appears reasonable. Indeed, calculations

of the interstitial dumbbell concentration for liquid aluminum using three independent methods give $c_i \approx 0.08$ [15], which agrees with Granato's estimate for copper ($c_i \approx 0.09$) [10]. On the other hand, computer simulations of aluminum melting have shown [23] that the shear modulus decreases from 14.9 GPa just below T_m to 1.8 GPa just above T_m. According to interstitialcy theory, the shear modulus of the melt is low but not zero. Using Equation (3), we estimate the interstitial dumbbell concentration at the melting temperature as:

$$c_i = \ln(14.9/1.8)/27 \approx 0.078,$$

which is close to the values obtained earlier. Finally, the pre-melting concentration of interstitial dumbbells in crystalline Al, based on precision shear modulus measurements, was found to be $c_i \approx 0.004$ [12]. The obtained estimate $c_i \approx 0.04$ after laser exposure is comparable to values at T_m, but, naturally, higher than pre-melting values.

Thus, these calculations are consistent with the understanding of the diaelastic effect in aluminum after laser exposure as a result of melting of the thin surface layer, accompanied by a sharp increase in the concentration of interstitial dumbbells, which remain largely frozen in the crystal due to high cooling rates. These frozen interstitial dumbbells define the observed diaelastic effect.

Other possible mechanisms for the shear modulus reduction after laser processing should also be considered. When exposed to a laser pulse, significant thermomechanical stresses arise in the sample. Assuming the temperature at the melt boundary is equal to the melting temperature of aluminum, while in the bulk of the sample, it remains close to room temperature, the temperature difference across the sample faces is approximately 600 K. Thus, the laser pulses generate high-amplitude mechanical pulses, which propagate through the entire sample. This corresponds to a relative total strain on the order of 10^{-2}. This is a fairly large value, which can lead to plastic deformation of the sample due to dislocation formation. As is well known, an increase in dislocation density can lead to a reduction in the shear modulus [24]. To evaluate this mechanism, it is necessary to assess the dislocation density in the samples after laser processing.

Moreover, at high laser intensities, the formation of shock waves may occur as a result of the breakdown of the ablation plume [25]. In this case, shock-wave-induced nanograin formation in the irradiated layer may take place, which could potentially contribute to the observed diaelastic effect.

A more detailed study of this phenomenon could provide new and important insights into the formation of the defect system in the crystal as a result of surface laser melting followed by high-speed cooling. It is also reasonable to expect that such experiments will lead to new significant information about the melting mechanism of simple metals.

5. CONCLUSION

For the first time, using resonant ultrasound spectroscopy (RUS), a diaelastic effect (reduction in elastic constants) has been detected in monocrystalline aluminum, induced by nanosecond ultraviolet laser pulses, which lead to melting of a thin near-surface layer of the sample. As a result of laser exposure, the shear modulus decreases from 0.87% to 1.45% with increasing incident energy density from 1.1 J/cm² to 5.3 J/cm². Thermal treatment by heating to the pre-melting temperature range restores the shear modulus to its initial values, while a significant increase in the amplitude of RUS peaks indicates a substantial reduction in the material's defect density.

The hypothesis is put forward that surface melting is accompanied by the formation of a high concentration of interstitial defects in a dumbbell configuration, which are preserved in the solid state due to the high cooling rate of the molten layer. The inelastic deformation caused by these defects leads to the observed diaelastic effect.

Other possible interpretations of this phenomenon are also noted.

ACKNOWLEDGMENTS

The authors express their gratitude to Professor D.L. Merson (Research Institute of Advanced Technologies, Tolyatti) for assistance in sample preparation, as well as to Professor I.A. Kaplunov, Associate Professor A.I. Ivanova (Tver State University), and Associate Professor V.V. Cheverikin (MISIS) for conducting electron microscopy studies.

FUNDING

The RUS measurements and analysis of results based on interstitial theory were supported by the Russian Science Foundation, Grant No. 23-12-00162.

The laser irradiation, structural studies, and part of the discussion were carried out with the support of the Russian Science Foundation, Grant No. 24-19-00727, using the resources of the Tver State University's TsKp.

REFERENCES

1. *C. B. Anisimov, Ya. A. Imas, G. S. Romanov, and Yu. V. Hodyko. Dejstvie izlucheniya bol'shoj moshchnosti na metally*, Nauka, Moscow (1970).
2. O. Zvelto. *Principy lazerov*, Lan', S-Pb, 2008.
3. *V. Yu. Homich, V. A. Shmakov*, Physics-Uspekhi **185**, 489 (2015).
4. *V. E. Fortov. Fizika vysokih plotnostej energii*, Fizmatlit, Moscow (2012).
5. *S. I. Ashitkov, S. A. Romashevskii, P. S. Komarov, A. A. Burmistrov, V. V. Zhakhovskii, N. A. Inogamov, and M. B. Agranat*, Quantum Electronics **45**, 547 (2015).
6. *V. Yu. Zheleznov, T. V. Malinskiy, S. I. Mikolutskiy, V. E. Rogalin, S. A. Filin, Yu. V. Khomich, V. A. Yamshchikov, I. A. Kaplunov, and A. I. Ivanova*, Pis'ma v ZhTF, **47**, 18 (2021).
7. *M. von Allmen, S. S. Lau, M. Maenpaa, B.-Y. Tsaur*, Appl. Phys. Lett. **36**, 205 (1980).
8. *L. Zhong, J. Wang, H. Sheng, Z. Zhang, and S. X. Mao*, Nature **512**, 177 (2014).
9. *Tong, X., Zhang, YE., Shang, BS. et al.* Nat. Mater. **23**, 1193 (2024)
10. *A. V. Granato*, Phys. Rev. Lett. **68**, 974 (1992).
11. *N. P. Kobelev, V. A. Khonik*, Physics-Uspekhi **193**, 717 (2023).
12. *E. V. Safonova, Yu. P. Mitrofanov, R. A. Konchakov, A. Yu. Vinogradov, N. P. Kobelev, V. A. Khonik*, J. Phys.: Condens. Matter. **28**, 1 (2016).
13. *E. V. Goncharova, A. S. Makarov, R. A. Konchakov, N. P. Kobelev, V. A. Khonik*, Pis'ma v ZhETF, **106**, 39 (2017).
14. *E. V. Safonova, R. A. Konchakov, Yu. P. Mitrofanov, N. P. Kobelev, A. Yu. Vinogradov, V. A. Khonik*, Pis'ma v ZhETF, **103**, 861 (2016)
15. *R. A. Konchakov, A. S. Makarov, A. S. Aronin, N. P. Kobelev, V. A. Khonik*, Pis'ma v ZhETF, **113** 341 (2021).
16. *J. M. Liu*, Optics Lett. **7**, 196 (1982).
17. *F. F. Balakirev, S. M. Ennaceur, R. J. Migliori, B. Maiorov, and A. Migliori*, Rev. Sci. Instrum. **90**, 121401 (2019).
18. *G. Simmons and H. Wang*, *Single Crystal Elastic Constants and Calculated Aggregate Properties*, Second edition, The MIT Press, Cambridge, MA (1971).
19. *A. V. Sokolov. Opticheskie svojstva metallov*, M.: Fizmatlit, 1961, 464 s.
20. *R. A. Konchakov, A. S. Makarov, G. V. Afonin, M. A. Kretova, N. P. Kobelev and V. A. Khonik*, Pis'ma v ZhETF **109**, 473 (2019).
21. *A. V. Granato*, Eur. J. Phys. B **87**, 18 (2014).
22. *C. A. Gordon and A. V. Granato*, Mater. Sci. Eng. A **370**, 83 (2004).
23. *M. Forsblom and G. Grimvall*, Nature Mater. **4**, 388 (2005).
24. *A. S. Nowick, B. S. Berry*, *Anelastic Relaxation in Crystalline Solids,* Academic Press, New York, London, (1972).
25. *A. A. Ionin, S. I. Kudryashov, and L. V. Seleznev*, Phys. Rev. E **82**, 016404 (2010).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 83–90

ORDER, DISORDER, AND PHASE TRANSITION IN CONDENSED MEDIA

TWO-DIMENSIONAL MAGNETOPLASMONS IN THE STRIP OF FINITE WIDTH

© 2025 R. Z. Vitlina[a], L. I. Magarill[a,b,*], A. V. Chaplik[a]

[a]Rzhanov Institute of Physics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia
[b]Novosibirsk State University, Novosibirsk, Russia
*e-mail: levm@isp.nsc.ru

Received June 10, 2024

Revised August 01, 2024

Accepted September 25, 2024

Abstract. Effect of sample boundaries on the spectrum of magnetoplasmons in the 2D electron gas was investigated. using the example of a strip. As should be expected in the limit of the plasmon wave length far exceeding the strip width the dispersion law of magnetoplasmons follows the one for 1D plasma waves however the leading term in the dispersion relation depends on the magnetic field. The dispersion laws of intraband plasmons in cases when one and two subbands are populated, depolarisation shift of the interband plasmon and spatial distribution of the plasmon electric field are found. The concentration and magnetic field dependencies of the plasmon frequency have been obtained numerically.

DOI: 10.31857/S00444510250108e6

1. INTRODUCTION

The edge magnetoplasmon (EMP) in a two-dimensional (2D) electron system was first theoretically studied in the works of Volkov and Mikhailov [1, 2]. The authors conducted both classical and quantum analyses for a half-plane and found the EMP dispersion relation as $\omega(k)$ where k is the 1D wave vector of the plasmon wave along the edge of the sample. Naturally, a question arises about the role of boundaries in a real experiment, particularly concerning plasma waves in a finite-width strip, where the influence of the opposite edge must also be taken into account. This formulation of the problem was outlined in the introduction of the paper by Balev and Vasilopoulos [3]. The authors proposed a strip model with "soft" walls, described by a parabolic potential for electrons near the strip boundaries. However, in their analysis of plasma oscillations, they effectively considered only one edge, naturally obtaining the already known result for the plasmon frequency. Meanwhile, the presence of the second boundary leads to qualitatively new features of the phenomenon: strictly speaking, one should not consider an edge plasmon but rather the eigenmodes of a planar plasma waveguide. It is important to note that in such a "waveguide", the electron motion is confined in one direction, while the electric field of the plasma wave extends formally to infinity. Within the framework of the classical hydrodynamic description of 2D plasma, this problem was solved in [4, 5]. The plasmon spectrum for a 2D electron strip under conditions of strong screening by a metallic electrode was found in [6], using a classical approach within the local capacitance approximation.

In the present study, we develop a quantum theory of magnetoplasmon waves in a 2D electron gas strip of finite width $L = 2w$. The boundary conditions for the wave functions correspond to hard walls, meaning the transverse electron motion (along the x-axis) corresponds to a "truncated" harmonic oscillator at $x = \pm w$ with the cyclotron frequency ω_c and a suspension point $X = -pl^2$, where p is the conserved y-component of the electron momentum in the Landau gauge, and l is the magnetic length ($\hbar = 1$). For the Landau level with index n, the wave function has the form:

$$\Psi_{n,X}(x,y) = N_{n,X}\varphi_{n,X}(x)\frac{\exp(ipy)}{\sqrt{L_y}}. \qquad (1)$$

Here $N_{n,X}$ is the normalization coefficient, and L_y is the length of the strip. For the wave function $\psi_{n,X}(x)$, we have (see, for example, [7]):

$$\psi_{n,X}(x) = e^{-(x-X)^2/2l^2} \times$$

$$\times\left[\Phi\left(-q_n(X)/2,1/2,(x-X)^2/l^2\right) - \right.$$

$$\left. -B(x-X)\Phi\left((1-q_n(X))/2,3/2,(x-X)^2/l^2\right)\right]. \quad (2)$$

The first index of the confluent hypergeometric function in Equation (2) determines the energy of the Landau subbands:

$$q_n(X) = E_n(X)/\omega_c - 1/2$$

via the dispersion equation following from the boundary conditions $\psi_{n,X}(x = \pm w) = 0$. From the same conditions, the constant B is determined.

The dispersion of the Landau subbands $E_n(X)$ is well known, and its graphs have been repeatedly presented in the literature in connection with studies of the quantum Hall effect (edge channels, edge states). The functions $\psi_{n,X}(x)$ and $E_n(X)$ are required to formulate the equation for plasma waves.

2. BASIC EQUATIONS

The problem considered here belongs to the class of plasma oscillations in multicomponent low-dimensional systems. The solution scheme, i.e., finding the eigenfrequencies of plasmons in such systems through the matrix dielectric function in the self-consistent field approximation, is described in [8] for 2D systems, such as quantum well structures with more than one populated transverse quantization level, double quantum wells, or multilayer superlattices.

In the case of magnetoplasmons in a 2D electron gas strip, the plasma components correspond to groups of electrons in different Landau levels (subbands $E_n(X)$), effectively forming 1D systems. Therefore, the Green's function of the Poisson equation takes the form of $G_k(x-x') = -K_0(|k(x-x')|)/2\pi$, where K_0 is the Macdonald function.

Another significant difference from [8] is the dependence of the transverse wave functions $\psi_{n,X}(x)$ (Equation (2)) on the longitudinal electron momentum p through the suspension point of the oscillator. Accounting for these distinctions, the equation for the matrix elements of the plasma wave potential $\varphi(x)e^{iky}$ takes the form (taking into account the selection rules for the momentum along the strip, which allow only transitions $(n,X) \to (m, X+kl^2)$:

$$\varphi_{n,X;m,X+kl^2} =$$

$$= \frac{2e^2}{\varepsilon L_y}\sum_{m',n',X'}\frac{f\left(E_{m'}(X'+kl^2)\right) - f\left(E_{n'}(X')\right)}{E_{m'}(X'+kl^2) - E_{n'}(X') + \omega + i\delta}\times$$

$$\times J_{m,n;m',n'}(X,X')\varphi_{n',X';m',X'+kl^2}, \quad (3)$$

Where ε is the average dielectric constant of the two media separated by the 2D electron gas, f is the Fermi occupation factor, and Form factors $J_{mn;m'n'}$ are defined as:

$$J_{m,n;m',n'}(X,X') =$$

$$= \int_{-w}^{w}\int_{-w}^{w} dxdx'\tilde{\psi}_{n,X}(x)\tilde{\psi}_{m,X+kl^2}(x)\times$$

$$\times K_0(|k(x-x')|)\tilde{\psi}_{n',X'}(x')\tilde{\psi}_{m',X'+kl^2}(x'). \quad (4)$$

In Equation (4), $\tilde{\psi}_{n,X}(x) = N_{n,X}\psi_{n,X}(x)$ represents the normalized wave function of the transverse motion. Thus, we obtain a system of linear homogeneous integral equations for the functions $\varphi_{n,X;m,X+kl^2}$, which we will denote by $\Phi_{nm}(X)$. For an unbounded discrete electron spectrum, the number of equations and, consequently, the number of different plasmon modes is infinite, even if only one level is populated, for example, $E_0(X)$. The off-diagonal terms in Equation (3) $m \neq n$ correspond to virtual transitions with an energy change of at least ω_c, i.e., they are responsible for inter-subband plasmons, whose spectrum has a gap $\Delta > \omega_c$ at zero wave vector $k = 0$. If one is interested only in the low-frequency part of the plasmon spectrum $\omega \ll \omega_c$, it is necessary to restrict consideration to intra-subband plasmons $m = n$ and additionally require the long-wavelength approximation $kl \ll 1$. In the following, we will consider both intra-subband plasmons, and inter-subband plasmons from the lower part of the spectrum, i.e., those associated with the levels $E_0(X)$ and $E_1(X)$.

3. INTRA-BAND PLASMON OF THE ZERO SUBBAND

In this case, instead of Equation (3), we have:

$$\Phi_{00}(X) =$$

$$= \frac{e^2}{\pi\varepsilon l^2}\int dX'\frac{f(E_0(X'+kl^2)) - f(E_0(X'))}{E_0(X'+kl^2) - E_0(X') + \omega + i\delta}\times$$

$$\times J_{00,00}(X,X')\Phi_{00}(X'). \quad (5)$$

Assuming $k \ll p \sim p_F$ (where p_F is the Fermi momentum), we expand the differences in Equation (5) up to the linear term in k. In the form factors J, we set k=0.

For $T = 0$, the numerator becomes $\delta(E_0(X') - E_F)$ (where E_F is the Fermi energy), and the integral reduces to the sum of two terms, corresponding to the values of the integrand at the points $X' = \pm X_0$, where $\pm X_0$ are the roots of the equation:

$E_0(X) = E_F$. Here ($E_0(X)$ is an even function of X.

By substituting variable X in left-hand side of Equation (5) with $\pm X_0$, we arrive at two linear homogeneous equations for the quantities $\Phi_\pm \equiv \Phi_{00}(\pm X_0)$:

$$\begin{aligned}\Phi_+ &= \beta k\left(\frac{J_{+-}}{\omega - kV_0}\Phi_- - \frac{J_{++}}{\omega + kV_0}\Phi_+\right),\\ \Phi_- &= \beta k\left(\frac{J_{--}}{\omega - kV_0}\Phi_- - \frac{J_{-+}}{\omega + kV_0}\Phi_+\right),\end{aligned} \tag{6}$$

Where $\beta = e^2 / \pi\varepsilon, V_0$ is Fermi velocity in the zero subband, while

$$J_{\pm\pm} = J_{00;00}(\pm X_0, \pm X_0),$$

$$J_{\pm\mp} = J_{00;00}(\pm X_0, \mp X_0).$$

It is evident that $J_{-+} = J_{+-}$. In the Appendix, it is shown that $J_{--} = J_{++}$. Thus, there are two independent form factors. The roots of the determinant of the system (6) determine the plasmon frequency $\omega_0(k)$:

$$\omega_0^2(k) = k^2\left(V_0^2 + \beta^2(J_{++}^2 - J_{+-}^2) + 2\beta V_0 J_{++}\right). \tag{7}$$

In the integrals defining $J_{+\pm}$, the functions $\psi_0^2(x)$ are localized near the points $\pm X_0$ within a region of order l. Therefore, for J_{++}, the argument K_0 is small under $k \to 0$, and we can use the asymptotic form of the Macdonald function:

$$K_0(|\, k(x - x')\,|) = -\ln(|\, k(x - x')\,|e^\gamma / 2),$$

where γ is the Euler constant. Then, for J_{++}, we obtain:

$$J_{++} = \ln\left(\frac{2e^{-\gamma}}{|\, k\,|\, l}\right) + \bar{J}_{++}, \tag{8}$$

where

$$\bar{J}_{++} = \int dx dx' \tilde{\psi}_{0,X_0}^2(x) \ln\left(\frac{l}{|\, x - x'\,|}\right)\tilde{\psi}_{0,X_0}^2(x'). \tag{9}$$

The leading term in J_{++} is $|\ln(|\, k\,|\, l)\,|$. For the form factor J_{+-}, the argument y of the K_0 function can be set to $2\,|\, k\,|\, X_0$, which may not be small, even for $kl \ll 1$. In this case $J_{+-} = K_0(2\,|\, k\,|\, X_0)$ and gives a significant contribution, provided the stronger condition $kX_0 \ll 1$ is satisfied. Under this condition, the plasmon frequency becomes

$$\omega_0^2(k) = k^2\left\{2\beta[\beta(\bar{J}_{++} - \bar{J}_{+-}) + V_0]\ln\left(\frac{2e^{-\gamma}}{|\, k\,|\, l}\right) + \right.$$

$$\left. + V_0^2 + \beta^2\left(\bar{J}_{++}^2 - \bar{J}_{+-}^2\right) + 2\beta V_0 \bar{J}_{++}\right\}. \tag{10}$$

Thus, we obtain the expected result for a one-dimensional (1D) plasmon, as found in [9, 10]:

$$\omega \sim k\sqrt{\left|\ln(|k|l)\right|}.$$

However, it is important to note that in the case considered here, the dependence of the magnetoplasmon frequency on the electron concentration and magnetic field cannot be expressed analytically. Another important difference is the change in the coefficient before the logarithmic term: to the Fermi velocity V_0 (for a 1D plasmon without a magnetic field), the first term in the square brackets of Equation (10) is added. This additional term can significantly exceed V_0 (for example, at $N_L = 10^6$ cm^{-1}, $H = 1.6$ T, the enhancement is more than an order of magnitude). The results of the numerical calculation are presented below.

The formulas derived in this section are valid up to the very beginning of the plasmon spectrum (k=0), when the plasmon wavelength is much larger than all characteristic lengths of the problem, including the width of the strip L. In this limit, the system effectively becomes one-dimensional. However, the transition to the half-plane limit, studied in [1, 2], is impossible, as it corresponds to an infinitely large width L. The dispersion laws differ: in the half-plane it is proportional to $\ln k$, while in stripe it is $\sqrt{\ln k}$ as expected for one-dimensional systems [9, 10].

4. INTRA-SUBBAND PLASMONS IN A TWO-SUBBAND SYSTEM

Let us now consider the case where the states $E_0(X)$ and $E_1(X)$ are populated, but we neglect the off-diagonal contribution $\Phi_{0,1}$. The Fermi level lies between $E_1(0)$ and $E_2(0)$, intersecting the curves

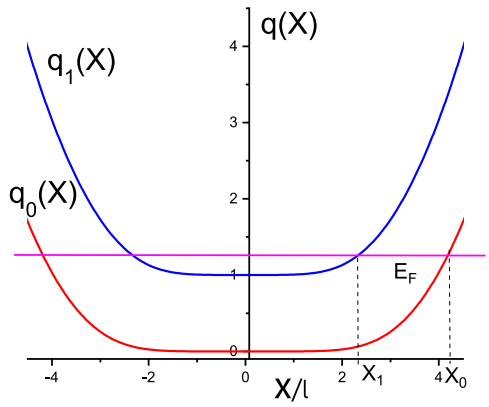

Fig. 1. Electron spectrum of the strip. The figure shows the two lowest Landau subbands. The horizontal line indicates the position of the Fermi level, $w/l = 4$.

$E_0(X)$ and $E_1(X)$ at the points $\pm X_0$ and $\pm X_1$, respectively (Fig. 1).

The four equations for $\Phi_{00}(\pm X_0)$ and $\Phi_{11}(\pm X_1)$ lead to a biquadratic equation for the plasmon frequencies, the roots of which are equal (here, the results are presented for infinitesimally small plasmon momenta $kX_0 \ll 1$, in order to clarify the behavior of $\omega(k)$ at the very beginning of the spectrum):

$$\omega_{ac}^2 = \frac{k^2}{2}\Big(V_0^2 + V_1^2 + \beta^2\Big[\bar{J}_{0+;0+}^2 - \bar{J}_{0+;0-}^2 + \bar{J}_{1+;1+}^2 - \bar{J}_{1+;1-}^2 + 2\bar{J}_{0+;1+}^2 - 2\bar{J}_{0+;1-}^2\Big] + 2\beta(V_0\bar{J}_{0+;0+} + V_1\bar{J}_{1+;1+})\Big), \quad (11)$$

$$\omega_{opt}^2 = \omega_{ac}^2 + 2k^2 \ln\left(\frac{2e^{-\gamma}}{|k|l}\right)\Big[\beta(V_0 + V_1) + \beta^2\Big(\bar{J}_{0+;0+} - \bar{J}_{0+;0-} + \bar{J}_{1+;1+} - \bar{J}_{1+;1-} + 2\bar{J}_{0+;1+} - 2\bar{J}_{0+;1-}\Big)\Big]. \quad (12)$$

Here, $V_{0,1}$ are the Fermi velocities in the zero and first subbands, respectively, while six independent form factors such as $\bar{J}_{0+,0+}, \bar{J}_{0+,1+}$, etc. are defined similarly to how it was done in the previous section.

It is important to emphasize that in Equation (11), all logarithmic contributions exactly cancel. The corresponding root of the dispersion equation gives the linear dependence $\omega_{ac}(k)$ as $k \to 0$, which justifies calling this branch acoustic. The second root (optical branch, Equation (12)) exhibits the known singularity at zero at $k \to 0$:

$$\omega_{opt}^2(k) \sim k^2\big|\ln(|k|l)\big|.$$

5. INTER-SUBBAND PLASMON IN A TWO-LEVEL SYSTEM

The rank of the characteristic determinant, considering N subbands, is N^2, since the dielectric function is a 4x4 matrix. Out of the N^2 roots, N correspond to intra-subband plasmons, while in the remaining $N(N-1)$ roots each pair gives rise to one inter-subband branch, making the total number of inter-subband branches equal to $N(N-1)/2$. We focus on the lowest inter-subband branch, associated with the E_0 and E_1 levels. The solution of the problem in the general case (for arbitrary plasmon momenta k) involves extremely complex numerical calculations, as neither the dispersion relations of electrons nor the form factors can be expressed analytically. Therefore, we limit ourselves to finding the threshold frequency $\omega_{01}(k=0)$, which determines the gap in the inter-subband plasmon spectrum. The difference between this value and the minimum energy gap between the E_0 and E_1 subbands is known as the depolarization shift.

If we retain only the equations for $m = 0,1$ and $n = 0,1$ in the system (3) and take the limit $k \to 0$, the right-hand side will only include the off-diagonal element φ, since the diagonal elements vanish due to the difference in occupation numbers approaching zero at $\psi_{0,X}(x)$, $\psi_{1,X}(x)$. In the same limit, the function $K_0(|k(x-x')|)$ simplifies to:

$$\ln(2e^{-\gamma}/|k(x-x')|) = \ln(2e^{-\gamma}/|k|l) + \ln(l/|x-x'|).$$

The first term does not contribute to the form factor $J_{01,01}$ due to the orthogonality of the wave functions $\psi_{0,X}(x)$, $\psi_{1,X}(x)$. As a result, we arrive at the equation

$$\Phi_{01}(X) = \frac{2\beta}{l^2}\int_{-X_0}^{X_0} dX' \times \frac{\Delta(X')}{\omega^2 - \Delta(X')^2} Q(X,X')\Phi_{01}(X'), \quad (13)$$

where $\Delta(X) = E_1(X) - E_0(X)$ and ω^2 is the desired eigenvalue (its minimum vakue is required, i.e. ω_{min}^2), and the kernel factor $Q(X,X')$ is equal to

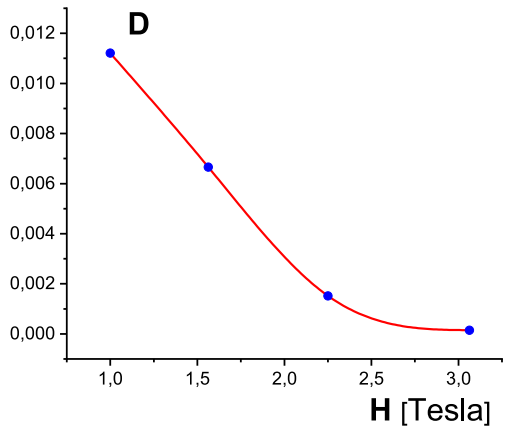

Fig. 2. Dependence of the depolarization shift of the inter-subband plasmon between levels 0 and 1 on the magnetic field; $D = \Omega/\Delta(X=0) - 1$, $N_L = 0.47 \cdot 10^6$ cm^{-1}, $L = 0.1$ μm.

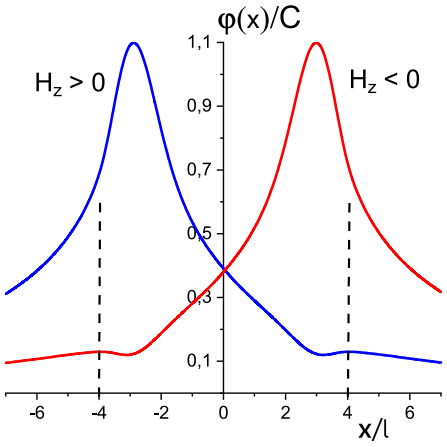

Fig. 3. Distribution of the plasmon wave potential across the transverse coordinate for two opposite propagation directions or magnetic field orientations; $N_L = 10^6$ cm^{-1}, $L = 0.2$ μm, $H = 1$ T.

$$Q(X,X') = \int_{-w}^{w}\int_{-w}^{w} dx dx' \tilde{\psi}_{0,X}(x)\tilde{\psi}_{1,X}(x) \times$$
$$\times \ln(l/|x-x'|)\tilde{\psi}_{0,X'}(x')\tilde{\psi}_{1,X'}(x'). \qquad (14)$$

The value of ω_{min}^2 was found numerically. We replaced the integral with the corresponding Riemann sum by dividing the integration interval into a large number of points, reducing the problem to finding the eigenvalues of a system of linear homogeneous equations, the number of which equals the number of partition points. The depolarization shift Ω is defined as the difference between the minimum plasmon frequency ω_{min} and the minimum energy gap between the levels Δ(0). Its dependence on the magnetic field is shown in Fig. 2.

As is known, the depolarization shift also determines the frequency of IR absorption during an inter-subband (inter-level in an infinite plane) transition, which differs from the energy gap due to the dynamic screening of the electric field of the exciting wave.

6. SPATIAL DISTRIBUTION OF THE PLASMON WAVE FIELD

In this section, we derive the expression for the coordinate dependence of the plasmon potential $\varphi(x)$, corresponding to the zeroth subband, i.e., the lowest-frequency branch of the plasmon spectrum. Within the self-consistent field theory, $\varphi(x)$ satisfies the Poisson equation (quasistatic approximation, neglecting retardation), with the right-hand side containing the electron density perturbation induced by the plasmon wave. In the present case, we consider only the contribution from the zeroth subband:

$$\Delta_{x,z}\varphi_0(x,z,k) - k^2\varphi_0(x,z,k) =$$
$$= -\frac{4\pi e^2}{\varepsilon L_y}\delta(z)\sum_X \frac{f\left(E_0(X+kl^2)\right) - f\left(E_0(X)\right)}{E_0(X+kl^2) - E_0(X) + \omega + i\delta} \times$$
$$\times \Phi_{00}(X)\tilde{\psi}_{0,X}^2(x). \qquad (15)$$

Equation (15) corresponds to a plasmon in the form of a plane wave Ce^{iky}, and the matrix element $\Phi_{00}(X)$ on the right-hand side is evaluated in the plane of the strip $z = 0$. The solution to Equation (15) is written using the Green's function $G(x-x')$, already defined in Section 2 for the plane $z = 0$. The resulting integral for $\varphi_0(x)$ in the long-wavelength limit and for $T = 0$ is evaluated similarly to the calculation of the plasmon frequency $\omega_0(k)$.

Now it is necessary to find the solutions of the system of two equations (6) for the matrix elements $\Phi_{00}(X)$ at the points $\pm X_0$. The result has the form (C is the wave amplitude determined by the excitation conditions):

$$\varphi_0(x) = Ck\beta\left(\frac{I_-(x)}{\omega_0(k) - kV_0} - \frac{RI_+(x)}{\omega_0(k) + kV_0}\right),$$
$$I_\pm(x) = \int_{-w}^{w} dx' K_0\left(|k(x-x')|\right)\tilde{\psi}_{0,\pm X_0}^2(x'), \qquad (16)$$
$$R = \frac{\omega_0(k) + kV_0}{\omega_0(k) - kV_0}\frac{k\beta J_{++} - \omega_0(k) + kV_0}{k\beta J_{+-}}.$$

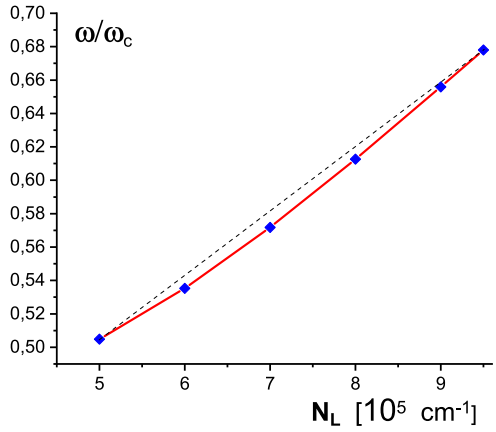

Fig. 4. Dependence of the plasmon frequency on the linear electron concentration. Magnetic field $H = 1$ T, strip width $L = 0.2$ μm.

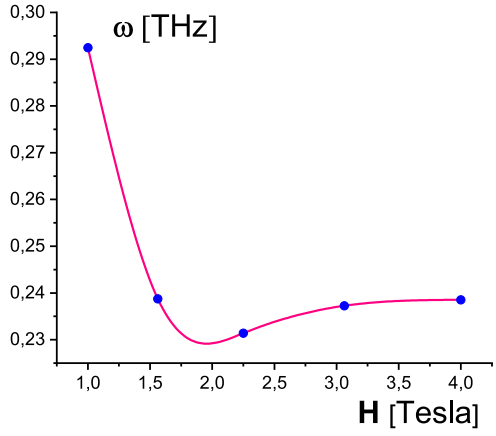

Fig. 5. Magnetic field dependence of the plasmon frequency; $N_L = 10^6$ cm^{-1}; $L = 0.2$ μm, $k = 0.4 \cdot 10^6$ cm^{-1}.

Fig. 3 shows the plasmon field $\varphi_0(x)$ for opposite propagation directions. As can be seen, for a given propagation direction, the maximum of $\varphi(x)$ is located near one edge of the strip. This result was previously obtained in [2] within the framework of the hydrodynamic approach.

The same mirror reflection occurs when the magnetic field direction is reversed: it is easy to see that under X, one should understand $-pl^2 sign(H)$, while $l^2 = c/|eH|$. Therefore, when the sign of H is changed, the points X_0 and $-X_0$ are swapped. This "reflection" of the plasmon field relative to the midline of the strip when the magnetic field sign is reversed is, in principle, accessible to experimental observation. When $l \ll w$ and the Fermi energy is such that the points $\pm X_0$ are close to the strip edges, the maximum of $\varphi(x)$ is also near one of the edges, and in this sense, such a wave can be called an edge magnetoplasmon.

7. DEPENDENCE ON CONCENTRATION AND MAGNETIC FIELD

The electron dispersion $E_0(p)$ (see Fig. 1) differs significantly from the standard parabolic law $p^2/2m$. Accordingly, all characteristics of the magnetoplasmon in the strip (the frequency dependence on electron concentration and magnetic field) appear unusual. For the intra-subband plasmon of the zeroth subband, the system is effectively one-dimensional, so $p_F = \pi N_L/2$, where N_L is the linear electron density (spin splitting is neglected), and $X_0 = \pi N_L l^2/2$. The dependence of E_F on V_0 is given by the right half of the lower curve in Fig. 1. The dependence of the plasmon frequency ω_0 on the linear density is determined by the Fermi velocity V_0 and the form factors X_0, appearing in formula (7). The results are presented in Fig. 4.

The dashed line in this figure is drawn to highlight the superlinear character of the dependence. Recall in this context that the classical 2D plasmon has a frequency that depends sublinearly on the surface density N_s:

$$\omega = (\omega_c^2 + \omega_p^2)^{1/2},$$

where $\omega_p^2 \propto N_s$.

The magnetic dispersion of the plasmon is even more unusual: the curve in Fig. 5 has a minimum at $H \approx 2$ T. This occurs because, as seen from (10), the dependence of the plasmon frequency on the magnetic field is due to two types of contributions. The terms containing the Fermi velocity V_0 provide the descending part of the curve in Fig. 5, as at a given density, the Fermi level rapidly decreases with increasing H and approaches the flat region of the electronic dispersion $E_0(p)$ where V_0 vanishes. Then the main contribution remains the first (Coulombic) term in (10), which leads to a logarithmically slow increase in the frequency.

For the depolarization shift (see Fig. 2), a rapid decrease is characteristic with a relatively small increase in H: more than an order of magnitude decrease at $\delta H / H = 75\%$. As the field increases, the behavior of the electron wave functions approaches

that realized in an infinite plane, as the influence of the strip boundaries decreases. However, in an infinite plane, $\Omega = 0$, because in a strong magnetic field, screening (at least linear screening) is absent, along with the electron density perturbations linear in the perturbing potential.

8. CONCLUSION

We have demonstrated that the consideration of sample boundaries significantly affects the magnetoplasmonic oscillations of a two-dimensional electron gas. Mathematically, the problem becomes considerably more complex due to the non-standard dispersion law of "magnetized" electrons – the dependence of energy on the conserved momentum component in the Landau gauge. In the simple case of a straight strip, it is possible to analytically obtain only the dispersion of intra-subband plasmons in the long-wavelength limit, corresponding to the lower part of the plasmon spectrum, which generally contains an infinite number of branches. The dependence of the plasmon frequency on concentration and magnetic field was determined using numerical methods.

ACKNOWLEDGMENTS

The authors are grateful to V.A. Volkov and I.V. Kukushkin for their assistance in reviewing the relevant bibliography and to A.V. Nenashev for helpful advice during the numerical calculations.

APPENDIX

Here we demonstrate the validity of the relationship $J_{--} = J_{++}$.. For this, we need the expression for $\psi_{n,X}(x)$, which already accounts for the boundary conditions. It has the form:

$$\psi_{n,X}(x) = e^{-(x-X)^2/2l^2} \times$$
$$\times\Big[\Phi\Big(-q_n(X)/2,1/2,(x-X)^2/l^2\Big) -$$
$$-\Phi\Big((1-q_n(X))/2,3/2,(x-X)^2/l^2\Big)\times$$
$$\times\frac{(x-X)\Phi\Big(-q_n(X)/2,1\,/\,2,(w-X)^2/l^2\Big)}{(w-X)\Phi\Big((1-q_n(X))/2,3/2,(w-X)^2/l^2\Big)}\Big]. \quad (17)$$

Furthermore, we need the explicit form of the equation defining the electron spectrum, i.e., the parameter $q_n(X)$. For it, we have:

$$F(q_n(X)) = 0, \quad (18)$$

$$F(q) = \frac{\Phi\Big(-q/2,1/2,(w-X)^2/l^2\Big)}{(w-X)\Phi\Big((1-q)/2,3/2,(w-X)^2/l^2\Big)} +$$
$$+\frac{\Phi\Big(-q/2,1/2,(w+X)^2/l^2\Big)}{(w+X)\Phi\Big((1-q)/2,3/2,(w+X)^2/l^2\Big)}. \quad (19)$$

Using the explicit expressions for the form factors $J_{\pm\pm}$, we write the difference $J_{--} - J_{++}$ as:

$$J_{--} - J_{++} = \int_{-w}^{w}\int_{-w}^{w} dxdx' K_0(|k(x-x')|)\times$$
$$\times\Big[N_{0,-X}^4\psi_{0,-X}^2(x)\psi_{0,-X}^2(x') -$$
$$-N_{0,X}^4\psi_{0,X}^2(x)\psi_{0,X}^2(x')\Big], \quad (20)$$

Here, $\psi_{n,X}(x)$ is defined in (17). By changing the integration variable in the first term within the square brackets, we arrive at the expression:

$$J_{--} - J_{++} = \int_{-w}^{w}\int_{-w}^{w} dxdx' K_0(|k(x-x')|)\times$$
$$\times\Big[N_{0,-X}^4\psi_{0,-X}^2(-x)\psi_{0,-X}^2(-x') -$$
$$-N_{0,X}^4\psi_{0,X}^2(x)\psi_{0,X}^2(x')\Big]. \quad (21)$$

It is evident that to prove the equality $J_{--} = J_{++}$, it is sufficient to show that the relationships $\psi_{0,-X}(-x) = \psi_{0,X}(x)$ and $N_{0,-X} = N_{0,X}$. hold. Using (17), we obtain:

$$\psi_{0,-X}(-x) - \psi_{0,X}(x) = e^{-(x-X)^2/2l^2}(x-X)\times$$
$$\times\Phi\Big((1-q_0(X))/2,3/2,(x-X)^2/l^2\Big)\times$$
$$\times\Bigg[\frac{\Phi\Big(-q_0(X)/2,1/2,(w-X)^2/l^2\Big)}{(w-X)\Phi\Big((1-q_0(X))/2,3/2,(w-X)^2/l^2\Big)} +$$
$$+\frac{\Phi\Big(-q_0(X)/2,1/2,(w+X)^2/l^2\Big)}{(w+X)\Phi\Big((1-q_0(X))/2,3/2,(w+X)^2/l^2\Big)}\Bigg]. \quad (22)$$

Thus, the expression under the square brackets in (22) is the function $F(q_0(X))$, defined in (19), and therefore:

$$\psi_{0,-X}(-x) = \psi_{0,X}(x). \quad (23)$$

For $N_{0,-X}$, we have:

$$N_{0,-X} = \left(\int_{-w}^{w} dx \psi_{0,-X}^2(x) \right)^{-1/2}.$$

Performing the variable change $x \to -x$ in the integral over x and considering (23), the evenness of the normalization coefficient with respect to X is thus proven.

REFERENCES

1. *V. A. Volkov, S. A. Mikhailov*, JETP Lett. **42**, 556 (1985) [Eksp. Teor. Fiz. Pis. Red. **42**, 450 (1985)].
2. *V. F. Volkov, S. A. Mikhailov*, Sov. Phys. JETP, **67**, 1639 (1988) [Zh. Eksp. Teor. Fiz **94**, 217 (1988)].
3. *O. G. Balev, V. Vasilopulos*, Phys. Rev. B **59**, 2807 (1999).
4. *I. L. Aleiner, D. Xue, and L. I. Glazman*, Phys. Rev. B **51**, 13467 (1995).
5. *G. Eliasson, J. Wu, P. Hawrylak, and J. J. Quinn*, Solid State Commun. **60**, 41 (1986).
6. *D. A. Rodionov, I. V. Zagorodnev*, JETP Letters **118**, 100 (2023) [Pis'ma Zh. Eksp. Teor. Fiz **118**, вып.2, 90 (2023)].
7. *A. H. MacDonald, P. Streda*, Phys. Rev.B **29**, 1616 (1984).
8. *R. Z. Vitlina, A. V. Chaplik*, Sov. Phys. JETP **54**, 536 (1981) [Zh. Eksp. Teor. Fiz **81**, 1011 (1981)].
9. *A. V. Chaplik and M. V. Krasheninnikov*, Surface Science **98**, 533 (1980).
10. *S. Das Sarma and W. Y. Lai*, Phys. Rev. B **32**, 1401 (1985).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 91–109

ORDER, DISORDER, AND PHASE TRANSITION IN CONDENSED MEDIA

QUANTUM *SU*(3)-FERRIMAGNET ON TRIANGULAR LATTICE IN MAGNETIC FIELD

© 2025 A. S. Martynov, D. M. Dzebisashvili*

Kirensky Institute of Physics, Federal Research Center KSC Siberian Branch, Russian Academy of Science, Krasnoyarsk, Russia

**e-mail: ddm@iph.krasn.ru*

Received July 21, 2024,

Revised August 30, 2024

Accepted September 04, 2024

Abstract. The phase diagrams (magnetic field H – single-ion anisotropy D) for three-sublattice $SU(3)$-ferrimagnet on triangular lattice with mixed sublattice spins ($S = 1,1/2,1/2$) at different values of exchange parameters I (between spins $S = 1$ and $S = 1/2$) and J (between spins $S = 1/2$) are calculated. To correctly account for the algebra of the $SU(3)$ group generators, which includes quadrupole operators, the representation of Hubbard operators was used. It is shown that depending on the system parameters there can be implemented ferrimagnetic Y- or inverted Y ($\bar{Y}$)-phase, canted V-phase (spins $S = 1/2$ are parallel), fan-shaped W-phase, as well as collinear ferrimagnetic and ferromagnetic phases. In the case of $I < J$, a line appears on the phase diagram on which $SU(3)$-ferrimagnet splits into two independent subsystems, one of which is paramagnetic with spins $S = 1$, and the second one is antiferromagnetic with spins $S = 1/2$ in a zero effective magnetic field. In the spin-wave approximation, the dependences of the average values of the quadrupole moment and dipole moments of the three sublattices on the magnetic field and the single-ion anisotropy are calculated. The spin-wave excitation spectrum is analyzed both at $I > J$ and at $I < J$. It is shown that at $I = J$ in the $SU(3)$-ferrimagnet, an accident degeneracy occurs, which can be lifted by taking into account quantum fluctuations.

Keywords: *mixed spin SU(3)-ferrimagnet, single-ion anisotropy, triangular lattice, phase diagrams in magnetic field*.

DOI: 10.31857/S00444510250109e9

1. INTRODUCTION

In recent years, there has been a significant increase in interest in materials where relativistic spin-orbit interaction leads to the manifestation of quantum effects on a macroscopic scale [1, 2]. These materials are commonly referred to as quantum magnets [3]. One of the most striking manifestations of quantum effects is the significant reduction of the average spin value in magnets with $S > 1/2$ [4]. The reason for the spin reduction lies in the consideration of single-ion anisotropy (SIA) arising from spin-orbit interaction or in the inclusion of pairwise interactions associated with higher-order spin invariants of the form $(S_f S_g)^{2S}$ [5–15]. In magnetic systems where such non-Heisenberg interactions are sufficiently strong, spin-nematic phases have been observed. These phases are characterized by zero magnetization even at zero temperature (i.e., complete spin reduction), but they exhibit spontaneous symmetry breaking due to quadrupole order parameters (mean values of operators bilinear in spin components) [10]. The enhancement of such quantum effects is facilitated by frustration [2], low temperature, low system dimensionality [16], and multi-sublattice structures.

For example, in multi-sublattice ferrimagnets with different magnetic ions, the manifestation of quantum effects can be significantly amplified due to the possible compensation of the effective field acting on the spins of magnetically active ions [17–26]. Indeed, as shown in [27], in a two-sublattice ferrimagnet, quantum spin reduction in the anisotropic sublattice (with $S = 1$) at low temperatures can be substantially suppressed by the exchange interaction field from the isotropic sublattice ($S = 1/2$). If there are more than two sublattices, the total effective field from two isotropically antiferromagnetically coupled sublattices acting on the ions of the third anisotropic

sublattice can be nullified, thereby eliminating the mentioned mechanism of spin reduction suppression.

In this regard, one of the key objectives in the theory of quantum magnets is to find a microscopic model that could predict and study new quantum effects with both experimental and practical significance. As outlined above, one promising approach is to investigate the combined action of multiple factors that promote quantum magnetism phenomena. In the context of this research direction, studies such as [28–30] proposed a model of a three-sublattice ferrimagnet with mixed spins $S = 1,1/2,1/2$ on a triangular lattice with Ising exchange interaction and SIA in the $S = 1$ spin subsystem. In those studies, based on Monte Carlo simulations, the main focus was on constructing phase diagrams in the temperature–SIA plane and searching for a technologically significant compensation regime, where the total magnetization reaches zero below the critical temperature. Notably, alongside SIA in the $S = 1$ spin subsystem, the model proposed in [28–30] possessed essential features such as low dimensionality and geometric frustration, which, as mentioned earlier, enhance quantum effects.

In a recent study [31], the authors investigated the SU(3) ferrimagnet (SU3F) model, which closely resembles the model proposed in [28–30] but includes two crucial generalizations. First, instead of Ising exchange interaction, the SU3F model employs isotropic Heisenberg exchange. It is well known that transverse components of exchange interaction in noncollinear magnetic structures induce zero-point quantum fluctuations, leading to antiferromagnetic (AF) fluctuations. These AF fluctuations, like SIA, can cause quantum spin reduction, and therefore, the quantum effects driven by AF and SIA should be distinguished. The second major difference between SU3F and the model proposed in [28–30] lies in the use of different exchange integrals I and J for interactions between the $S = 1$ and $S = 1/2$ sublattices and between the two $S = 1/2$ sublattices, respectively. As shown below, the phase diagrams of SU3F differ qualitatively depending on the ratio between the exchange integrals.

Furthermore, it is essential to highlight an important conceptual feature of the SU3F model. This feature is associated with the fact that significant SIA, as known from previous studies [8–15, 32–36], necessitates the inclusion of the full set of generators of the SU(3) algebra acting in the Hilbert space of the $S = 1$ spin states. Therefore, conventional spin operators are insufficient for describing such systems. To emphasize this aspect, the model proposed in [31] was named the quantum SU(3) ferrimagnet model.

The general characteristic of the SU3F model is the simultaneous consideration of several factors that enhance quantum effects: SIA, AF fluctuations, multi-sublattice structure, low dimensionality, and exchange frustration.

The study of the SU3F model in [31] was conducted in the absence of an external magnetic field and at zero temperature. The dependence of the sublattice spin moments and the quadrupole moment on the SIA parameter was calculated for different exchange integral ratios I/J. It was found that the critical value of the SIA parameter D_c, at which SU3F transitions to the quadrupole phase, can be significantly smaller than both I and J. Moreover, for $I > J$, a compensation point was observed in the total moment M dependence on the SIA parameter, i.e., M at $D < D_c$.

This work represents a logical continuation of the studies conducted in [31]. Its primary goal is to construct the phase diagram of SU3F in the external magnetic field–SIA parameter plane and to analyze the modification of the magnetic structure and order parameters when crossing the phase boundaries. The ground state energy and the corresponding spin configuration are calculated within the mean-field approximation at zero temperature. This condition, as is well known, is unachievable by the Monte Carlo method used in the previously cited works [28–30]. To correctly account for the SU(3) algebra generators in the $S = 1$ spin subsystem, the Hubbard operator formalism is employed [11, 35, 37]. In the calculation of order parameters, spin operator bosonization is applied: the Holstein–Primakoff transformation for the $S = 1/2$ spin subsystem and the indefinite metric formalism for the $S = 1$ subsystem [11, 14].

The remainder of this paper is organized as follows. Section 2 formulates the SU3F Hamiltonian in an external magnetic field lying in the easy-plane direction. Section 3 presents the $SU(2)$ transformation of the S = 1/2S = 1/2S = 1/2 spin operators, corresponding to the rotation of local coordinate axes. Section 4 details the Holstein–Primakoff transformation for the $S = 1/2$ spin subsystem. Section 5 describes the transition to the Hubbard operator representation and their triple

$SU(3)$ transformation for diagonalizing the single-ion Hamiltonian of the $S = 1$ spin subsystem. The bosonization of Hubbard operators and the derivation of the dispersion equation are covered in Section 6. Sections 7 and 8 analyze the characteristics of phase diagrams and the changes in order parameters for $I < J$ and $I > J$, respectively. Section 9 demonstrates the degeneracy of the SU3F mean-field ground state at $I = J$. Section 10 discusses changes in the spin-wave excitation spectrum as the magnetic field increases under different exchange parameter ratios. The main conclusions of the study are presented in Section 11.

2. MODEL OF $SU3$-FERRIMAGNETISM

The crystal structure of the considered SU3F is shown in Fig. 1. The red circles mark the lattice sites of the sublattice with spin value $S = 1$, further referred to as the L-sublattice. The green and blue circles mark the lattice sites of the sublattices with spin value $S = 1/2$, denoted further as F and G sublattices, respectively. The periodicity of the system is defined by the basis vectors a_1 and a_2, equal in magnitude. The vectors z and x connect the nodes of different sublattices.

The Hamiltonian of SU3F in an external magnetic field can be written as:

$$\mathcal{H} = \mathcal{H}_A + \mathcal{H}_{exch} + \mathcal{H}_{field}, \tag{1}$$

where:

$$\mathcal{H}_{exch} = J\sum_{\{fg\}} S_f S_g + I\sum_{\{fl\}} S_f S_l + I\sum_{\{gl\}} S_g S_l,$$

$$\mathcal{H}_A = D\sum_l \left(S_l^y\right)^2, \tag{2}$$

$$\mathcal{H}_{field} = -h\sum_f S_f^z - h\sum_g S_g^z - h_L\sum_l S_l^z.$$

The operator $\mathcal{H}_{exch}$ describes the pairwise exchange interaction between the nearest-neighbor spins from different sublattices. The lower indices f, g and l of the spin operators denote the lattice sites from the F-, G- and L sublattices, respectively. The exchange integral J determines the strength of the antiferromagnetic interaction between the nearest-neighbor spins from the F- and G- sublattices, while the integral I governs the interaction between the $F(G)$- and L- sublattices. The curly brackets under the summation symbols in (2) indicate that the summation is carried out only over nearest neighbors, with each pair of nodes counted only once.

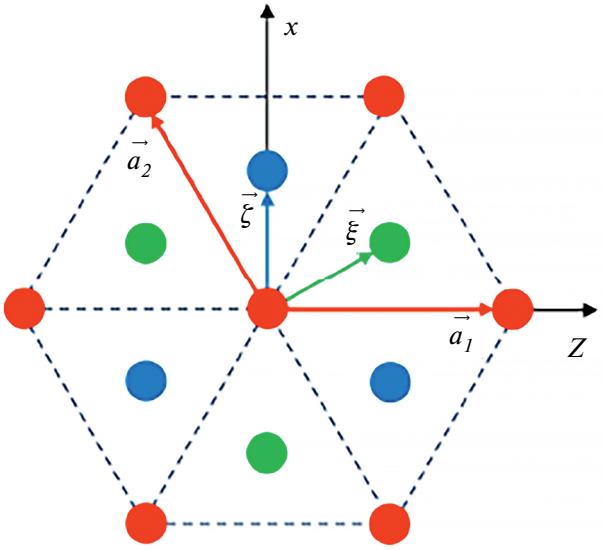

Fig. 1. Crystal structure of the three-sublattice SU3F on a triangular lattice. Red, green, and blue circles indicate the positions of the nodes in the L-, F- and G- sublattices, respectively. $|a_1| = |a_2| = a$ are the Bravais lattice vectors, while ξ and ζ represent the basis vectors.

The operator $\mathcal{H}_A$ describes the effect of single-ion anisotropy (SIA) of the easy-plane type acting on the spins $S = 1$ in the L- sublattice. The anisotropy parameter D is positive. The y axis is directed perpendicular to the ferrimagnet plane xz, which is, therefore, the easy magnetization plane.

The operator $\mathcal{H}_{field}$ accounts for the Zeeman energy of the spins in the external magnetic field H, lying in the ferrimagnet plane (easy plane) and determining the parameters $h = g\propto_B H$, and $h_L = g_L\propto_B H$, where $\propto_B$ is the Bohr magneton, and g and g_L are the Landé factors for the $F(G)$-sublattices, respectively. In general, the g-factors may differ for different sublattices. In this study, we assume that the moments are formed without the participation of orbital degrees of freedom, i.e., they are purely spin-related, so $g_L = g = 2$.

The direction of the magnetic field and the type of SIA ensure that the average moment of the L sublattice of R_L, is oriented in the xz plane, perpendicular to the anisotropy axis y. Furthermore, considering the nature of the exchange interactions and the results of Ref. [38], it can be argued that the magnetic structure of the SU3F ground state for any values of D and H is characterized by a planar

configuration of spin expectation values. Therefore, without loss of generality, we will assume that the spins of all three sublattices lie in the ferrimagnet plane xz, with the z axis of the original coordinate system conveniently directed along the magnetic field.

3. $SU(3)$-TRANSFORMATION OF THE HAMILTONIAN

To calculate the ground-state energy of SU3F, it is convenient to start with a unitary transformation of the Hamiltonian H:

$$\mathcal{H}(\theta_F,\theta_G) = U_2(\theta_F,\theta_G)\mathcal{H}U_2^{+}(\theta_F,\theta_G), \qquad (3)$$

with the operator

$$U_2(\theta) = \prod_{f\in F}\exp\left(-i\theta_F S_f^y\right)\prod_{g\in G}\exp\left(-i\theta_G S_g^y\right). \qquad (4)$$

The transformation (3) allows one to switch to new local coordinates for the F - and G - sublattices, where the quantization axes z' and z'' are rotated by the angles θ_F and θ_G around the y axis, aligning them along the equilibrium magnetizations R_F and R_G, respectively (see Fig. 2).

The unitary transformation (3) of the Hamiltonian (1) corresponds to the following formal substitution of the spin operators for the F- and G- sublattices [39]:

$$S_f^x \to S_f^x\cos\theta_F + S_f^z\sin\theta_F,\ S_f^y \to S_f^y,$$
$$S_f^z \to S_f^z\cos\theta_F - S_f^x\sin\theta_F, \qquad (5)$$
$$S_g^x \to S_g^x\cos\theta_G + S_g^z\sin\theta_G,\ S_g^y \to S_g^y,$$
$$S_g^z \to S_g^z\cos\theta_G - S_g^x\sin\theta_G. \qquad (6)$$

As a result, the Hamiltonian operator (1) is transformed into the following form:

$$\mathcal{H} = D\sum_l (S_l^y)^2 +$$
$$+J\sum_{\{fg\}}\left\{(S_f^xS_g^x + S_f^zS_g^z)\cos(\theta_F-\theta_G) + \right.$$
$$\left.+S_f^yS_g^y + (S_f^zS_g^x - S_f^xS_g^z)\sin(\theta_F-\theta_G)\right\} +$$
$$+I\sum_{\{fl\}}\left\{(S_f^xS_l^x + S_f^zS_l^z)\cos\theta_F + S_f^yS_l^y + \right.$$
$$\left.+(S_f^zS_l^x - S_f^xS_l^z)\sin\theta_F\right\} +$$

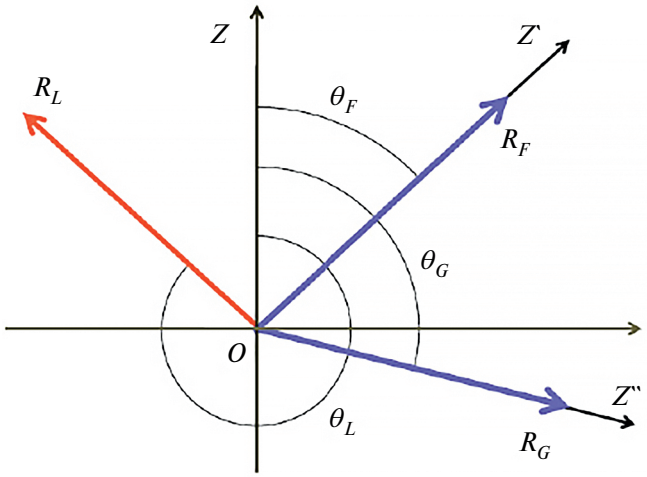

Fig. 2. Rotation of local coordinate axes during the unitary transformation (3). In the F- and G- sublattices with $S = 1/2$, the axes z are rotated by the angles θ_F and θ_G, taking new positions z' and z'', respectively. The local coordinates in the L- subsystem with $S = 1$ remain unchanged, while the angle formed by the moment $R_L$and the z axis is denoted by θ_L.

$$+I\sum_{\{gl\}}\left\{(S_g^xS_l^x + S_g^zS_l^z)\cos\theta_G + S_g^yS_l^y + \right.$$
$$\left.+(S_g^zS_l^x - S_g^xS_l^z)\sin\theta_G\right\} -$$
$$-h\sum_f\left\{S_f^z\cos\theta_F - S_f^x\sin\theta_F\right\} -$$
$$-h\sum_g\left\{S_g^z\cos\theta_G - S_g^x\sin\theta_G\right\} - h_L\sum_l S_l^z, \qquad (7)$$

where the operators S_f^β and S_g^β ($\beta = x,y,z$) relating to F- and G-subsystems refer to the projections of the spin moments on the quantization axes corresponding to index β in the new (rotated) local coordinate systems.

4. HOLSTEIN–PRIMAKOFF TRANSFORMATION

Following the strategy outlined in the introduction for calculating the ground-state energy of SU3F, we perform the Holstein–Primakoff transformation separately for the F- and G- sublattices:

$$S_f^+ = \sqrt{2S - a_f^+a_f}\cdot a_f,\quad S_f^z = S - a_f^+a_f,$$
$$S_g^+ = \sqrt{2S - b_g^+b_g}\cdot b_g,\quad S_f^z = S - b_g^+b_g, \qquad (8)$$

where the bosonic creation $a_f^+(b_g^+)$ and annihilation $a_f(b_g)$ operators describe spin transitions at site $f(g)$ of the $F(G)$ sublattice from the state $|\uparrow'\rangle(|\uparrow''\rangle)$, corresponding to spin orientation along the $z'(z'')$

axis, to the opposite orientation $|\downarrow'\rangle(|\downarrow''\rangle)$, and vice versa.

Substituting (8) into the Hamiltonian (7) gives the result:

$$\mathcal{H} = E_0 + \mathcal{H}^{(0)} + \mathcal{H}^{(1)} + \mathcal{H}^{(2)}. \tag{9}$$

where:

$$E_0 = J_0 S^2 N \cos(\theta_F - \theta_G) - \\ -hSN(\cos\theta_F + \cos\theta_G), \tag{10}$$

Here, $H^{(n)}$ $(n = 0,1,2)$ denotes the number of sites in the sublattice. The Hamiltonian $H^{(0)}$ represents the sum of the single-ion Hamiltonians for the L- subsystem:

$$\mathcal{H}^{(0)} = \sum_l \mathcal{H}_0(l),$$

where:

$$\mathcal{H}_0(l) = D(S_l^y)^2 + \bar{H}_z S_l^z + \bar{H}_x S_l^x, \tag{11}$$

The effective fields are defined as:

$$\bar{H}_z = I_0 S(\cos\theta_F + \cos\theta_G) - h_L, \\ \bar{H}_x = I_0 S(\sin\theta_F + \sin\theta_G), \; I_0 = 3I. \tag{12}$$

The linear term in bosonic operators from the Hamiltonian (9) can be written as:

$$\mathcal{H}^{(1)} = \sum_{\{fl\}} I\sqrt{\frac{S}{2}}\left[\cos\theta_F S_l^x - \sin\theta_F S_l^z\right](a_f + a_f^+) + \\ + \sum_f \sqrt{\frac{S}{2}}\left[J_0 S\sin(\theta_G - \theta_F) + h\sin\theta_F\right](a_f + a_f^+) + \\ + \sum_{\{gl\}} I\sqrt{\frac{S}{2}}\left[\cos\theta_G S_l^x - \sin\theta_G S_l^z\right](b_g + b_g^+) + \\ + \sum_g \sqrt{\frac{S}{2}}\left[J_0 S\sin(\theta_F - \theta_G) + h\sin\theta_G\right](b_g + b_g^+) + \\ + \frac{I}{i}\sqrt{\frac{S}{2}}\left\{\sum_{\{fl\}} S_l^y (a_f - a_f^+) + \sum_{\{gl\}} S_l^y (b_g - b_g^+)\right\}, \tag{13}$$

where:

$$J_0 = 3J$$

The last term in expression (9) describes the excitations in the F- and G subsystems and has the form:

$$\mathcal{H}^{(2)} = J\frac{S}{2}\sum_{\{f,g\}}\left\{\left[(a_f + a_f^+)(b_g + b_g^+) - \right.\right. \\ \left. - 2(a_f^+ a_f + b_g^+ b_g)\right]\cos(\theta_F - \theta_G) - \\ - (a_f - a_f^+)(b_g - b_g^+)\Big\} - \\ - I\sum_{\{f,l\}}(\cos\theta_F S_l^z + \sin\theta_F S_l^x)a_f^+ a_f - \\ - I\sum_{\{g,l\}}(\cos\theta_G S_l^z + \sin\theta_G S_l^x)b_g^+ b_g + \\ + h\cos\theta_F\sum_f a_f^+ a_f + h\cos\theta_G\sum_g b_g^+ b_g. \tag{14}$$

Next, the mean-field logic dictates replacing the spin operators $H^{(1)}$ and $H^{(2)}$ of the L- subsystems with their average values. In the considered zero-temperature regime, averaging the operators S_l^α $(\alpha = x,y,z)$ is sufficient to perform based on the ground state of the single-site Hamiltonian (11).

5. DIAGONALIZATION OF THE SINGLE-ION HAMILTONIAN

To diagonalize the single-ion Hamiltonian (11), we use the approach developed in [40]. We transition from spin operators to Hubbard operators [37], where $X_l^{m,n} = |m\rangle\langle n|$, where $m,n = \{-1,0,+1\}$ are the eigenstates of the operator S_l^z with corresponding eigenvalues $|m\rangle$ and $|n\rangle$ of $S_l^z|n\rangle = n|n\rangle$. Substituting

$$S_l^x = \frac{1}{\sqrt{2}}\left(X_l^{1,0} + X_l^{\bar{1},0} + X_l^{0,1} + X_l^{0,\bar{1}}\right), \\ S_l^y = \frac{i}{\sqrt{2}}\left(-X_l^{1,0} + X_l^{\bar{1},0} + X_l^{0,1} - X_l^{0,\bar{1}}\right), \\ (S_l^y)^2 = \frac{1}{2}\left(X_l^{1,\bar{1}} + X_l^{\bar{1},1} - X_l^{1,1} - X_l^{\bar{1},\bar{1}}\right) + X_l^{0,0}, \\ S_l^z = X_l^{1,1} - X_l^{\bar{1},\bar{1}}, \quad \bar{1} \equiv -1, \tag{15}$$

Describing the transition to representation of the Hubbard operators into the single-ion Hamiltonian (11) gives:

$$\mathcal{H}_0(l) = \left(\frac{D}{2} + \bar{H}_z\right)X_l^{1,1} + DX_l^{0,0} + \\ + \left(\frac{D}{2} - \bar{H}_z\right)X_l^{\bar{1},\bar{1}} - \frac{D}{2}\left(X_l^{1,\bar{1}} + X_l^{\bar{1},1}\right) + \\ + \frac{\bar{H}_x}{\sqrt{2}}\left(X_l^{1,0} + X_l^{0,1} + X_l^{\bar{1},0} + X_l^{0,\bar{1}}\right). \tag{16}$$

In the absence of a magnetic field, the ground state of the system is degenerate with respect to rotations around the y axis. Choosing the x axis

along the vector R_L and assuming equivalence between the F - and G -sublattices ($\theta_F = -\theta_G$,), the parameter φ\phiφ becomes zero, and the last term in (16) vanishes. In this case, the Hamiltonian $H_0(l)$ couples only two of the three states ($|+1\rangle$ and $|-1\rangle$), and its diagonalization requires only one unitary transformation (see [31]). However, the presence of a magnetic field couples all three states $|n\rangle$ ($n = \{-1,0,+1\}$), requiring three consecutive transformations for the diagonalization of the single-ion Hamiltonian.

The unitary operator $U_{nm}(\alpha,l)$ for each transformation is defined by its generator $\Gamma_{nm}(l) = X_l^{nm} - X_l^{mn}$ from the SU(3) group, according to the expression:

$$U_{nm}(\alpha,l) = \exp\{\alpha\Gamma_{nm}(l)\} =$$
$$= 1 + (\cos\alpha - 1)(X_l^{nm} + X_l^{mn}) + \sin\alpha\,\Gamma_{nm}(l). \quad (17)$$

The new Hubbard operators $X_l^{\tilde{r}\tilde{s}} = |\tilde{r},l\rangle\langle\tilde{s},l|$, defined through the new basis states

$$|\tilde{r},l\rangle = U_{nm}(-\alpha,l)|r,l\rangle, \quad (18)$$

are expressed via original Hubbard operators as follows:

$$X_l^{\tilde{r}\tilde{s}} = U_{nm}(-\alpha,l)\,X_l^{rs}\,U_{nm}^{+}(-\alpha,l). \quad (19)$$

Thus, the unitary transformation reduces to a simple substitution in the single-site Hamiltonian:

$$X_l^{rs} \to U_{\tilde{n}\tilde{m}}(\alpha,l)\,X_l^{\tilde{r}\tilde{s}}\,U_{\tilde{n}\tilde{m}}^{+}(\alpha,l). \quad (20)$$

Explicit expressions for the right-hand side of the last formula were derived in [40] and are provided in Appendix A for completeness. The variational parameter α in (17) is chosen such that the off-diagonal terms $X_l^{\tilde{n}\tilde{m}}$ and $X_l^{\tilde{m}\tilde{n}}$ vanish in the transformed Hamiltonian.

Performing the three consecutive unitary transformations with the operators $U_{1,0}(\alpha_2)$, $U_{0,-1}(\alpha_3)$ and $U_{1,-1}(\alpha_1)$, following the rule (20), and retaining the original notation for the indices of the new states $n = \{-1,0,+1\}$ (i.e., without tildes), we obtain the diagonal form of the single-ion Hamiltonian $H_0(l)$:

$$\mathcal{H}_0(l) = \sum_n \varepsilon_n X_l^{nn}, \quad n = -1,0,+1. \quad (21)$$

The eigenvalues ε_n of the single-ion Hamiltonian can be expressed as ($\bar{1} = -1$):

$$\varepsilon_1 = e_{\bar{1},\bar{1}}\sin^2\alpha_1 + e_{1,1}\cos^2\alpha_1 + e_{1,\bar{1}}\sin 2\alpha_1,$$
$$\varepsilon_{\bar{1}} = e_{\bar{1},\bar{1}}\cos^2\alpha_1 + e_{1,1}\sin^2\alpha_1 - e_{1,\bar{1}}\sin 2\alpha_1, \quad (22)$$
$$\varepsilon_0 = e_{0,0},$$

where

$$e_{1,1} = D\sin^2\alpha_2 + \left(\frac{D}{2} + \bar{H}_z\right)\cos^2\alpha_2 +$$
$$+\frac{\bar{H}_x}{\sqrt{2}}\sin 2\alpha_2,$$
$$e_{\bar{1},\bar{1}} = D\cos^2\alpha_2\sin^2\alpha_3 - \frac{D}{2}\sin\alpha_2\sin 2\alpha_3 +$$
$$+\left(\frac{D}{2} + \bar{H}_z\right)\sin^2\alpha_2\sin^2\alpha_3 +$$
$$+\left(\frac{D}{2} - \bar{H}_z\right)\cos^2\alpha_3 - \quad (23)$$
$$-\frac{\bar{H}_x}{\sqrt{2}}(\cos\alpha_2\sin 2\alpha_3 + \sin 2\alpha_2\sin^2\alpha_3),$$
$$e_{0,0} = D\cos^2\alpha_2\cos^2\alpha_3 + \frac{D}{2}\sin\alpha_2\sin 2\alpha_3 +$$
$$+\left(\frac{D}{2}+\bar{H}_z\right)\sin^2\alpha_2\cos^2\alpha_3+\left(\frac{D}{2}-\bar{H}_z\right)\sin^2\alpha_3 +$$
$$+\frac{\bar{H}_x}{\sqrt{2}}(\cos\alpha_2\sin 2\alpha_3 - \sin 2\alpha_2\cos^2\alpha_3),$$

$$e_{1,\bar{1}} = \left(\frac{\bar{H}_z}{2} - \frac{D}{4}\right)\sin(2\alpha_2)\sin\alpha_3 -$$
$$-\frac{D}{2}\cos\alpha_2\cos\alpha_3 +$$
$$+\frac{\bar{H}_x}{\sqrt{2}}(-\cos 2\alpha_2\sin\alpha_3 + \sin\alpha_2\cos\alpha_3).$$

From the requirement of nullifying the coefficients of the non-diagonal X-operators in the transformed Hamiltonian, the following system of equations for the angles α_j ($j = 1,2,3$) is obtained:

$$\mathrm{tg}\alpha_3 = \frac{\left(\frac{D}{2} - \bar{H}_z\right)\sin 2\alpha_2 + \sqrt{2}\,\bar{H}_x\cos 2\alpha_2}{D\cos\alpha_2 - \sqrt{2}\,\bar{H}_x\sin\alpha_2},$$
$$\mathrm{tg}2\alpha_3 = \frac{\sqrt{2}\,\bar{H}_x\cos\alpha_2 + D\sin\alpha_2}{2\bar{H}_z + \left(\frac{D}{2} - \bar{H}_z\right)\cos^2\alpha_2 - \frac{\bar{H}_x}{\sqrt{2}}\sin 2\alpha_2}, \quad (24)$$
$$2\alpha_1 = 2e_{1,\bar{1}} / (e_{1,1} - e_{\bar{1},\bar{1}}).$$

Similarly, by sequentially applying formula (20) with the operators $U_{1,0}(\alpha_2)$, $U_{0,\bar{1}}(\alpha_3)$ and $U_{1,\bar{1}}(\alpha_1)$, to the representation (15), the spin operators S_l^x, S_l^y, S_l^z and $(S_l^y)^2$ can be expressed through the new (transformed) X-operators. The expansion coefficients of the spin operators S_l^α in terms of the new Hubbard operators X_l^{nm} will represent the matrix elements of the spin operators in the new states: $s_{n,m}^\alpha \equiv \langle n \mid S_l^\alpha \mid m \rangle$ ($\alpha = x,y,z$). Explicit expressions for these matrix elements are given in Appendix B.

Within the mean-field approximation, the spin operators in the Hamiltonian $H^{(1)}$ should be replaced by their average values, i.e., the diagonal matrix elements $s_{n,n}^\alpha$, calculated for the ground state $|n\rangle$, corresponding to the minimum value of ε_n. Below, we will choose the set of solutions of equations (24) for the angles α_j ($j = 1,2,3$) such that the state $|+1\rangle$ is the ground state.

Since $s_{nn}^y = 0$ for any n (see Appendix B), the last two sums in formula (13) for $H^{(1)}$ vanish. The reduction of the remaining terms in (13) occurs under the conditions:

$$\begin{aligned} &I_0(s_{1,1}^x \cos\theta_F - s_{1,1}^z \sin\theta_F) + \\ &+ J_0 S \sin(\theta_G - \theta_F) + h \sin\theta_F = 0, \\ &I_0(s_{1,1}^x \cos\theta_G - s_{1,1}^z \sin\theta_G) + \\ &+ J_0 S \sin(\theta_F - \theta_G) + h \sin\theta_G = 0, \end{aligned} \tag{25}$$

These conditions will be further used to determine the equilibrium values of the angles θ_F and θ_G. The angle θ_L, introduced in Fig. 2 for clarity, is not a tuning parameter and can be determined through the ratio of the average values of the spin projections S_l^z and S_l^x.

The magnetic structure of the SU3F ground state is determined by the solutions of the five equations (24) and (25) for the angles α_j ($j = 1,2,3$), θ_F and θ_G, followed by the selection of the solution set that corresponds to the minimum value of the mean-field energy of the entire system:

$$E_{MF} = E_0 + N\varepsilon_1, \tag{26}$$

where the values E_0 and ε_1 are defined by equations (10) and (22), respectively. In Section 7, the SU3F phase diagrams in the $h - D$ -coordinates, calculated based on the methodology presented here, will be presented.

6. BOSONIZATION OF THE L- SUBSYSTEM AND THE DISPERSION EQUATION

Within the chosen approximation, the ground state energy E_{MF} is determined without considering AF (antiferromagnetic) contributions. Therefore, the contributions from the last term in the Hamiltonian (9), quadratic in Bose operators, are absent in expression (26) for E_{MF}. Nevertheless, when calculating the dependencies of order parameters on the magnetic field and single-ion anisotropy (SIA), the energy spectrum of spin-wave excitations is required, and to determine this spectrum, the operator $H^{(2)}$ must be taken into account.

To compute the energy spectrum within the spin-wave approximation, we first express the spin operators through the new (transformed) X-operators. Using (15) and the formulas from Appendix A, we obtain expressions for the S-operators of the form:

$$S_l^\alpha = \sum_{n,m} s_{nm}^\alpha X_l^{nm}, \quad \alpha = x,y,z, \tag{27}$$

where the matrix elements s_{nm}^α are given in Appendix B.

Next, considering that the state spectrum $H_0(l)$ is characterized by three levels and the ground state of the single-ion Hamiltonian is the state $|+1\rangle$, we introduce, following [11, 14], two types of Bose operators: c and d. The creation of one $c(d)$ boson at site l is described by the creation operator $c_l^+(d_l^+)$ and corresponds to the system transitioning from the "vacuum" state $|+1\rangle$ to the state $|0\rangle(|-1\rangle)$ with one $c(d)$ boson. The Hermitian conjugate operator $c_l(d_l)$, acting in the opposite direction, annihilates the $c(d)$ boson. States with more than one boson are excluded by the metric operator as non-physical.

The representation of Hubbard operators through Bose operators, proposed in [40] within the framework of the indefinite metric formalism [41], takes the form:

$$\begin{aligned} &X_l^{1,0} = (1 - c_l^+ c_l - d_l^+ d_l) c_l,\ X_l^{0,1} = c_l^+, \\ &X_l^{1,\bar{1}} = (1 - c_l^+ c_l - d_l^+ d_l) d_l,\ X_l^{\bar{1},1} = d_l^+, \\ &X_l^{0,\bar{1}} = c_l^+ d_l,\ \ X_l^{\bar{1},0} = d_l^+ c_l,\ \ X_l^{0,0} = c_l^+ c_l, \\ &X_l^{\bar{1},\bar{1}} = d_l^+ d_l,\ \ X_l^{1,1} = (1 - c_l^+ c_l - d_l^+ d_l). \end{aligned} \tag{28}$$

We use the representation (28) in the formulas (27) and substitute the resulting expressions for the S-operators (see Appendix C) into the terms $H^{(1)}$

and $H^{(2)}$ of the Hamiltonian (9). As a result, an expression arises in which only contributions up to the second order in the a, b, c and d - operators should be retained. Performing the Fourier transform:

$$a_f = \frac{1}{\sqrt{N}}\sum_k e^{ikf} a_k, \quad b_g = \frac{1}{\sqrt{N}}\sum_k e^{ikg} b_k,$$
$$c_l = \frac{1}{\sqrt{N}}\sum_k e^{ikl} c_k, \quad d_l = \frac{1}{\sqrt{N}}\sum_k e^{ikl} d_k, \tag{29}$$

we obtain the desired Hamiltonian, which can be written as follows:

$$H = E_{MF} + H_{SW}. \tag{30}$$

Here, the first term E_{MF} corresponds to the ground state energy in the mean-field approximation (see formula (26)), while the second term H_{SW} describes spin-wave excitations and is defined by the expression:

$$\begin{aligned} H_{SW} = \sum_k \{ & E_a a_k^+ a_k + E_b b_k^+ b_k + E_c c_k^+ c_k + E_d d_k^+ d_k + \\ & + J_+(\gamma_k a_k^+ b_k + \gamma_k^* b_k^+ a_k)\} + \\ & + J_-(\gamma_k a_k^+ b_{-k}^+ + \gamma_k^* a_k b_{-k}) + \\ & + I_{0F}^+(\gamma_k c_k^+ a_k + \gamma_k^* a_k^+ c_k) + \\ & + I_{0F}^-(\gamma_k c_k^+ a_{-k}^+ + \gamma_k^* c_k a_{-k}) + \\ & + I_{\bar{1}F}^+(\gamma_k d_k^+ a_k + \gamma_k^* a_k^+ d_k) + \\ & + I_{\bar{1}F}^-(\gamma_k d_k^+ a_{-k}^+ + \gamma_k^* d_k a_{-k}) + \\ & + I_{0G}^+(\gamma_k^* c_k^+ b_k + \gamma_k b_k^+ c_k) + \\ & + I_{0G}^-(\gamma_k^* c_k^+ b_{-k}^+ + \gamma_k c_k b_{-k}) + \\ & + I_{\bar{1}G}^+(\gamma_k^* d_k^+ b_k + \gamma_k b_k^+ d_k) + \\ & + I_{\bar{1}G}^-(\gamma_k^* d_k^+ b_{-k}^+ + \gamma_k d_k b_{-k}). \end{aligned} \tag{31}$$

In this expression, the following notations were introduced:

$$\begin{aligned} E_a &= -J_0 S\cos(\theta_F - \theta_G) + h\cos\theta_F - \\ &\quad - I_0(s_{11}^z\cos\theta_F + s_{11}^x\sin\theta_F), \\ E_b &= -J_0 S\cos(\theta_G - \theta_F) + h\cos\theta_G - \\ &\quad - I_0(s_{11}^z\cos\theta_G + s_{11}^x\sin\theta_G), \end{aligned}$$

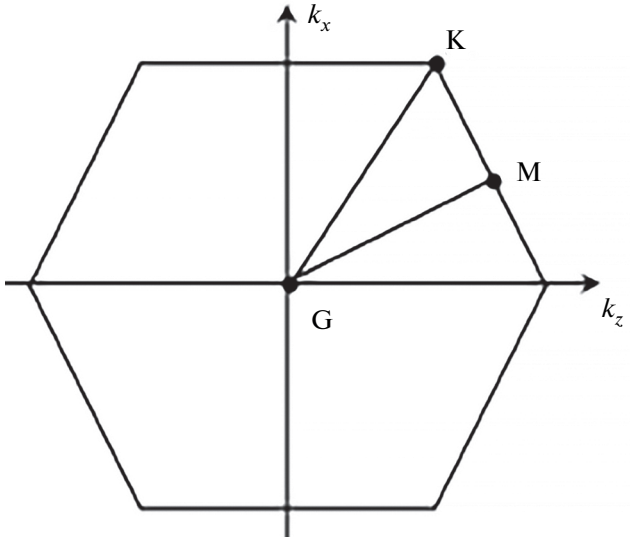

Fig. 3. The Brillouin zone of the triangular lattice and three high-symmetry points: Γ, *K*, *M*.

$$\begin{aligned} E_c &= \varepsilon_0 - \varepsilon_1, \quad E_d = \varepsilon_{\bar{1}} - \varepsilon_1, \\ J_\pm &= \frac{J_0 S}{2}(\cos(\theta_F - \theta_G) \pm 1), \\ I_{nA}^\pm &= I_0\sqrt{\frac{S}{2}}\left(s_{n1}^x\cos\theta_A - s_{n1}^z\sin\theta_A \pm \frac{s_{n1}^y}{i}\right), \\ n &= \{0,\bar{1}\}, \quad A = \{F,G\}, \\ \gamma_k &= \frac{1}{3}\sum_\delta e^{ik\delta} = \frac{1}{3}\left(2\cos\frac{k_z}{2}e^{i\frac{k_x}{2\sqrt{3}}} + e^{-i\frac{k_x}{\sqrt{3}}}\right). \end{aligned} \tag{32}$$

In the sum defining the triangular lattice invariant γ_k, the vector δ takes three values: $\{\xi, -\zeta, \zeta - \xi\}$ (see Fig. 1). The Brillouin zone, which bounds the region of quasimomentum values, is shown in Fig. 3.

To obtain the dispersion equation, we define the matrix retarded Green's function $\langle\langle X_k \mid X_k^+\rangle\rangle_\omega$,, where

$$X_k^+ = (a_k^+, b_k^+, c_k^+, d_k^+, a_{-k}, b_{-k}, c_{-k}, d_{-k}).$$

From the requirement for nontrivial solutions of the equation of motion for $\langle\langle X_k \mid X_k^+\rangle\rangle_\omega$, the spectrum equation follows:

$$\begin{vmatrix} \omega - A_k & -B_k \\ B_k & \omega + A_k \end{vmatrix} = 0, \tag{33}$$

where

$$A_k = \begin{pmatrix} E_a & J_+\gamma_k & I^+_{0F}\gamma^*_k & I^+_{\bar{1}F}\gamma^*_k \\ J_+\gamma^*_k & E_b & I^+_{0G}\gamma_k & I^+_{\bar{1}G}\gamma_k \\ I^+_{0F}\gamma_k & I^+_{0G}\gamma^*_k & E_c & 0 \\ I^+_{\bar{1}F}\gamma_k & I^+_{\bar{1}G}\gamma^*_k & 0 & E_d \end{pmatrix} \quad (34)$$

and

$$B_k = \begin{pmatrix} 0 & J_-\gamma_k & I^-_{0F}\gamma^*_k & I^-_{\bar{1}F}\gamma^*_k \\ J_-\gamma^*_k & 0 & I^-_{0G}\gamma_k & I^-_{\bar{1}G}\gamma_k \\ I^-_{0F}\gamma_k & I^-_{0G}\gamma^*_k & 0 & 0 \\ I^-_{\bar{1}F}\gamma_k & I^-_{\bar{1}G}\gamma^*_k & 0 & 0 \end{pmatrix}. \quad (35)$$

7. SU3F PHASE DIAGRAM FOR $I < J$

We will discuss the SU3F phase diagram in the magnetic field–anisotropy parameter (D) coordinates separately for three cases of exchange parameter ratios: $I < J$, $I > J$, $I = J$. In this section, we consider the first case: $I < J$.

Fig. 4 shows the phase diagram of the SU3F ground state, calculated according to the methodology outlined in Section 5, for the exchange parameter ratio $I / J = 0.8$. It is evident that three phases are realized in the considered regime: the inverted Y-phase (hereinafter referred to as $\bar{Y}$), the W-phase, and the ferromagnetic phase.

In the $\bar{Y}$ phase, the average spin vector of the L-sublattice, R_L is aligned along the magnetic field direction (the z-axis), while the average spin vectors of the F- and G- sublattices, R_F and R_G, form equal but opposite angles with the z-axis: $\theta_F = -\theta_G$. The magnitude of the angles θ_F and θ_G varies within the range $[\pi/2, \pi]$.

In the symmetric W-phase, the angles θ_F and θ_G also have equal magnitudes and opposite signs. However, unlike the $\bar{Y}$ phase, the range of these angle magnitudes is different: $[0,\pi/2]$. In this case, the projections of all three vectors R_F, R_G and R_L onto the z-axis are positive. The boundary between the $\bar{Y}$- and W-phases in Fig. 4 is marked by the dashed line. To the right of the red line on the phase diagram, the ferromagnetic phase is realized: the average spin vectors of the L-, F- and G- sublattices are aligned along the magnetic field.

The evolution of the magnetic structure as the magnetic field at $I < J$ is characterized by a monotonic decrease in the absolute values of

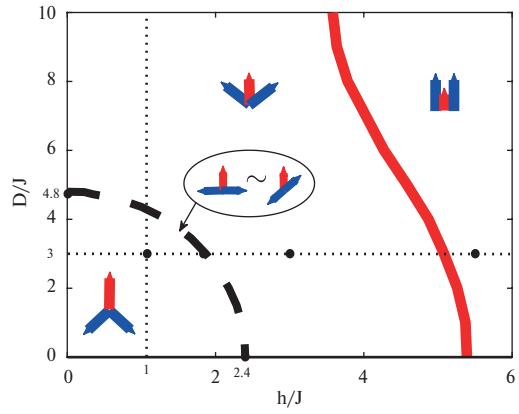

Fig. 4. The h–D phase diagram of the SU3F ground state for $I/J = 0.8$. The black dashed line corresponds to the boundary between the $\bar{Y}$- and W-phases, while the solid red line indicates the boundary between the W phase and the ferromagnetic phase. The pictograms illustrating the magnetic structure of SU3F represent the R_L vector with a red arrow, the $R_{F(G)}$, vectors with blue arrows, and the magnetic field h directed upward. On the dashed line, a phase is realized where the subsystems with spins S = 1 and 1/2 become effectively independent.

the angles θ_F and θ_G, as H increases which vanish at a certain field value dependent on the anisotropy parameter OA (see the red line in Fig. 4). This behavior is illustrated by three pictograms schematically depicting the magnetic structure in each of the three regions of the phase diagram.

To further understand the presented phase diagram, we will calculate the dependence of the SU3F order parameters on the magnetic field for a fixed anisotropy parameter OA and on the anisotropy parameter OA for a fixed magnetic field h.

The average spin values R_F and R_G in F- and G-sublattices can be calculated using the Holstein–Primakoff representation (8), according to which:

$$\begin{aligned} R_F &= \langle S^{z'}_f \rangle = S - n_a, \\ R_G &= \langle S^{z''}_g \rangle = S - n_b, \end{aligned} \quad (36)$$

where the boson occupation numbers $n_a = \langle a^+_f a_f \rangle$ and $n_b = \langle b^+_g b_g \rangle$ are computed using the spectral theorem from the matrix Green's function $\langle\langle X_k \mid X^+_k \rangle\rangle_\omega$ introduced in Section 6.

The average spin magnetic moment of the L-sublattice R_L can be found using the formula:

$$R_L = \sqrt{\left(R^z_L\right)^2 + \left(R^x_L\right)^2}, \quad (37)$$

where the quantities R^z_L and R^x_L are determined by the average occupation numbers of c - and d

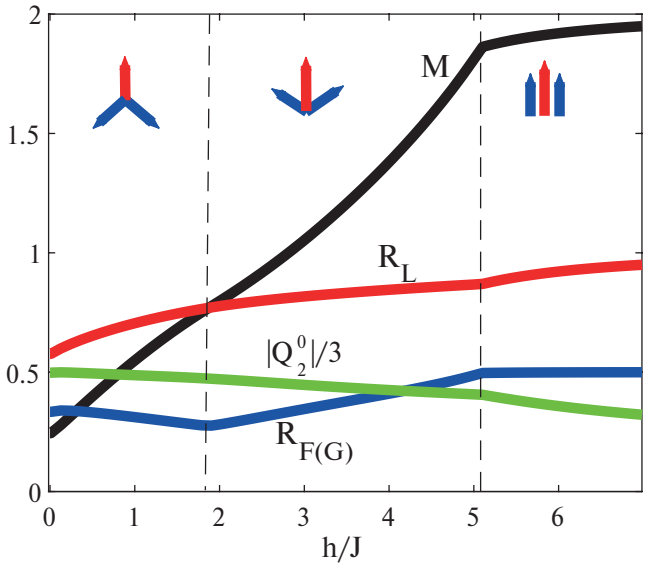

Fig. 5. Dependence of the total moment R_L (red line), $R_{F(G)}$ (blue line), M (black line), and $|Q_2^0|/3$ (green line) on the magnetic field h. The ratio between the exchange integrals is $I/J = 0.8$, while $D/J = 3$. The three pictograms composed of one red and two blue arrows have the same meaning as in Fig. 4.

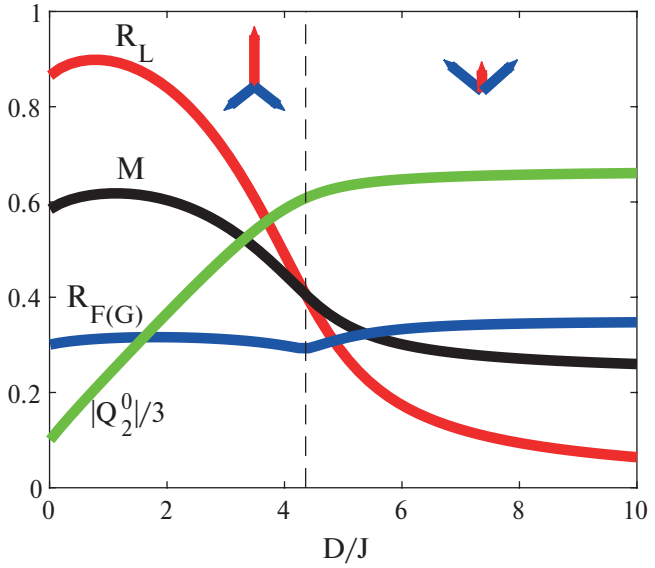

Fig. 6. Dependencies of the quantities R_L (red line), $R_{F(G)}$ (blue line), M (black line), and $|Q_2^0|/3$ (green line) on the anisotropy parameter OA. The ratio between the exchange integrals is $I/J = 0.8$, and $h/J = 1$

-bosons: $n_c = \langle c_k^+ c_k \rangle$ and $n_d = \langle d_k^+ d_k \rangle$, as well as correlators $\langle c_k^+ d_k \rangle$ and $\langle d_k^+ c_k \rangle$. The corresponding expressions are obtained by averaging the formulas given in Appendix C.

Since the total magnetic moment $M = R_F + R_G + R_L$ is directed along the external magnetic field (i.e., along the z-axis), its transverse component must identically vanish:

$$R_L^x + R_F \sin\theta_F + R_G \sin\theta_G = 0,$$

and the longitudinal component equals:

$$M = R_L^z + R_F \cos\theta_F + R_G \cos\theta_G . \qquad (38)$$

The average value of the quadrupole moment [42]:

$$Q_2^0(l) = 3\left(S_l^y\right)^2 - 2 \qquad (39)$$

is calculated similarly after averaging the corresponding formulas from Appendix C.

Fig. 5 shows the dependence of the total moment, the average spin magnetic moments R_L, $R_{F(G)}$, M (black line), and the quadrupole moment $|Q_2^0|/3$ on the magnetic field h for the anisotropy parameter OA $D/J = 3$ and the exchange integral ratio $I/J = 0.8$.

The change in the magnetic field on this figure corresponds to the movement along the horizontal dashed line on the phase diagram in Fig. 4. It is evident that at the transition point from the W phase to the ferromagnetic phase, all curves in Fig. 5 exhibit a kink. The values of M and R_L increase as expected with increasing field h, while the quadrupole moment decreases.

The dependencies of the order parameters M, R_L, R_F, R_G and Q_2^0 on the anisotropy parameter $h/J = 1$ are shown in Fig. 6.

The change in the anisotropy parameter D in this figure corresponds to the movement along the vertical dashed line on the phase diagram in Fig. 4. It can be seen that when crossing the boundary between the $\bar{Y}$ - and W -phases, the dependencies of the order parameters on D exhibit a kink, while the quadrupole moment saturates. The average moment of L-sublattice rapidly decreases near the phase boundary but decreases more slowly with further increase in D.

Clearly, the reduction of R_L facilitates the upward reorientation of the $R_{F(G)}$, vectors, as it reduces the exchange energy loss between the spins $S = 1$ and $S = 1/2$.

An important feature of the phase diagram presented in Fig. 4 is that along the entire boundary between the $\bar{Y}$ - and W -phases (black dashed line), the angle between the vectors R_F and R_G equals π. In this case, from expressions (12) for the effective fields, we find:

$$\bar{H}_z = -h_L, \quad \bar{H}_x = 0. \tag{40}$$

Taking into account these relations and the condition $h_L \neq 0$, the solutions of equations (24) for the angles α_j ($j = 1,2,3$) take the form:

$$2\alpha_1 = \frac{D}{2h_L}(-1)^{n+m}, \quad \alpha_2 = \pi n, \quad \alpha_3 = \pi m, \tag{41}$$

where n and m are integers.

Substituting these solutions into the expressions for the matrix elements of the spin operators from Appendix B gives:

$$s^z_{11} = \cos 2\alpha_1, \quad s^x_{11} = 0. \tag{42}$$

Since $s^x_{11} = 0$ and $\theta_F - \theta_G = \pi$, from equations (25) for the angles θ_F and θ_G, we find the condition:

$$s^z_{11} = h/I_0, \tag{43}$$

which must be satisfied by the matrix element s^z_{11} at the boundary between the $\bar{Y}$ - and W -phases. The equation describing the boundary of these phases can be easily obtained from the compatibility condition of the three equations for the angle α_1 and the matrix element s^z_{11} in formulas (41), (42), and (43).

As a result, the following relationship between the model parameters and the magnetic field is obtained:

$$D = \frac{2g_L}{g}\sqrt{I_0^2 - h^2}. \tag{44}$$

This expression analytically describes the dashed line in Fig. 4.

It is important to note that at the points of the phase diagram lying on this dashed line, the orientation of the (antiparallel) vectors R_F and R_G relative to the z-axis is not fixed. This fact implies the degeneracy of the SU3F ground state with respect to the simultaneous rotation of the spins from the F- and G- sublattices around the z-axis, provided that the vectors R_F and R_G remain antiparallel.

Indeed, substituting the solutions (41) for the angles α_j ($j = 1,2,3$) into formulas (22) and (23), as well as fixing the difference in π in exprssion (10) between angles θ_F and θ_G ,, we obtain:

$$\varepsilon_1 = D/2 - \sqrt{h_L^2 + (D/2)^2}, \quad E_0 = -J_0 S^2 N.$$

Thus, at the points of the phase diagram lying strictly on the boundary between the $\bar{Y}$ - and W- phases (i.e., along the dashed line in Fig. 4), the ground-state energy $E_{MF} = E_0 + N\varepsilon_1$ (see equation (26)) does not depend on the angles θ_F and θ_G.

The physical reason for this behavior is that, at $\theta_F - \theta_G = \pi$, the two effective fields acting on the spins in the L-sublattice from the F- and G-subsystems compensate each other (see equation (12)). As a result, the L-sublattice effectively "decouples" from both the F- and G-subsystems. Meanwhile, the external magnetic field h_L continues to act on the L-subsystem, aligning the vector R_L along the direction h_L.

Simultaneously, the F- and G-sublattices also "lose connection" with the L-subsystem, as the effective fields generated by it in the F- and G-sublattices are fully canceled by the external magnetic field h. Indeed, as follows from expression (7), the quantities E_a and E_b (see (32)) are precisely the effective fields acting on the spins in the F- and G-sublattices, respectively. Since, at the points lying on the dashed line of the phase diagram in Fig. 4, the conditions (42) and (43) are satisfied, the contributions to effective fields E_a and E_b from the L- subsystem ($-I_0 s^z_{11}\cos\theta_{F(G)}$), the external magnetic field ($h\cos\theta_{F(G)}$) cancel each other out.

Thus, at the points belonging to the dashed line on the phase diagram in Fig. 4, the SU3F system decouples into two effectively non-interacting subsystems: one formed by the S=1 spins of L-sublattice and the other by the $S = 1/2$ spins of F- and G-sublattice spins. In this case, the S=1 spins behave like a paramagnet in an external magnetic field, as they continue to experience the field h_L, while the interaction between them vanishes. The $S = 1/2$ spins behave like a two-sublattice (F and G) collinear antiferromagnet in an effective zero magnetic field. This condition, allowing for an arbitrary orientation of the antiferromagnetic vector in the zx -plane, leads to additional degeneracy of the ground state.

8. PHASE DIAGRAM OF SU3F AT $I/J > 1$

When $I > J$, the phase diagram of SU3F under a magnetic field changes qualitatively. Fig. 7 presents the phase diagram calculated for the exchange parameter ratio $I/J = 1.2$. It is evident that four magnetic phases are realized in this case: the Y-phase, the collinear ferrimagnetic phase, the V ($\bar{V}$) phase, and the ferromagnetic phase.

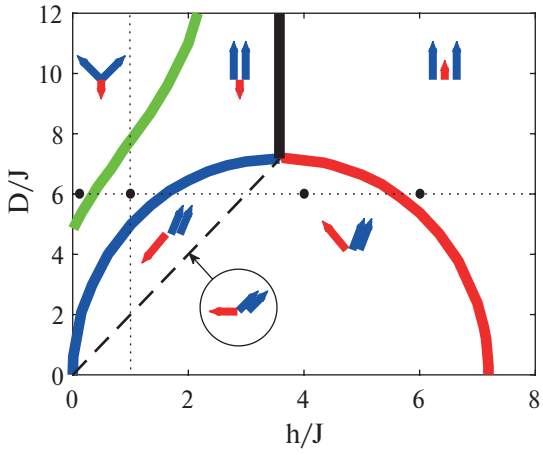

Fig. 7. Phase diagram of the SU3F ground state at $I/J = 1.2$. The green line denotes the boundary between the Y-phase and the collinear ferrimagnetic phase, the blue line separates the collinear ferrimagnetic and $\bar{V}$phases, the black line separates the ferromagnetic and collinear ferrimagnetic phases, the red line separates the ferromagnetic and M phases, and the dashed line marks the boundary between the $\bar{V}$- and V- phases (on this line, $\theta_L = -\pi/2$).

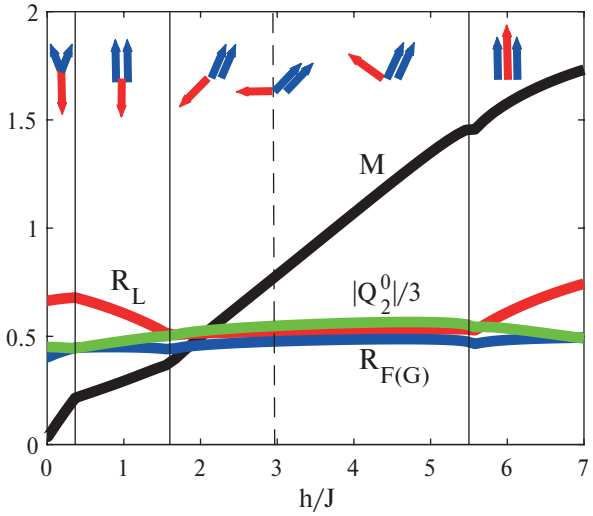

Fig. 8. Dependence of the quantities R_L (red line), $R_{F(G)}$ (blue line), M (black line), and $|Q_2^0|$ (green line) on the external magnetic field h at $I/J = 1.2$ and $D/J = 6$.

In the Y phase, the vector R_L, representing the average spin in the L-sublattice (red arrow in the pictograms of Fig. 7), is directed opposite to the magnetic field (along the $-z$ axis), while the average spin vectors R_F and R_G (blue arrows) in F- and G-sublattices form equal but opposite angles with the z-axis: z: $\theta_F = -\theta_G$, while $|\theta_{F(G)}| \in [0, \pi/2]$.

When transitioning from the Y phase to the collinear ferrimagnetic phase, the angles θ_F and θ_G simultaneously become zero, and all three vectors R_F, R_G and R_L become collinear: the first two align with the magnetic field, while the third opposes it.

Beneath the blue and red curves in Fig. 7 lies the so-called V phase, where the vector R_L forms a nonzero angle θ_L with the z-axis, while the vectors R_F and R_G form equal angles θ_F and θ_G. These angles vary within the range $0, \pi/2$.

This region can be further divided by a dashed line (shown in Fig. 7) into two subregions. To the right of this line, $|\theta_L| < \pi/2$, while to the left, $|\theta_L| > \pi/2$. We retain the V designation for the first region and label the second as the $\bar{V}$ phase for distinction. Along the entire dashed line, the angle θ_L strictly equals $\pi/2$.

In the ferromagnetic phase, all three vectors R_F, R_G and R_L align with the magnetic field.

As in the previous section, to understand the magnetic structure, we examine the changes in order parameters along two directions on the phase diagram: along the horizontal dashed line at a fixed value $D/J = 6$ and along the vertical dashed line at a fixed field $h/J = 1$ (see Fig. 7).

Fig. 8 shows the dependence of the quantities R_L, $R_{F(G)}$, M and Q_2^0 on the external magnetic field h at $D/J = 6$. This corresponds to movement along the horizontal dashed line in Fig. 7.

It is evident that the changes in $R_{F(G)}$ and Q_2^0 with increasing field h are minor, and the reduction in the average spin value $R_{F(G)}$ due to AF interactions is insignificant. In contrast, the average moment of L-sublattice is significantly suppressed due to both AF and OA interactions. In the ferrimagnetic phase, the vector R_L is directed opposite to the field, and its magnitude decreases with increasing h, as expected. In the ferromagnetic phase, the vector R_L aligns with the field, causing its magnitude to increase.

A crucial observation from the graphs in Fig. 8 is that the evolution of the magnetic structure follows the same sequence as in a triangular-lattice antiferromagnet (TLAF) with $S = 1/2$, but without OA [38, 43]. However, while the extended ferrimagnetic (or uud) phase in TLAF can only be explained by quantum fluctuations (which lift accidental degeneracy), in SU3F, this phase arises solely due to OA. Moreover, the behavior of the total moment M qualitatively reproduces the key stages of the TLAF's evolution: the monotonic increase of M in the Y-, $\bar{V}$- and V-phases; a plateau-like region in the ferrimagnetic (uud) phase (commonly referred

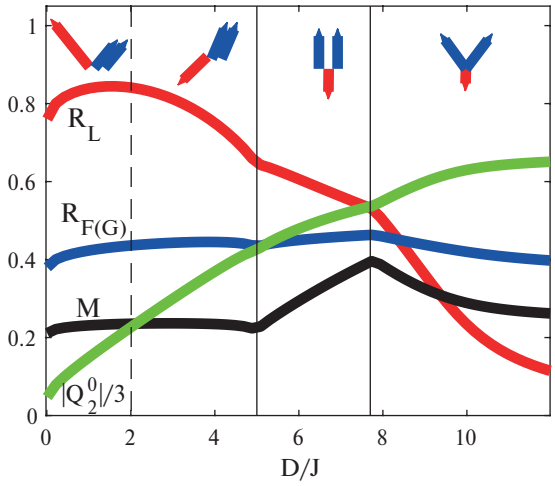

Fig. 9. Dependence of the quantities R_L (red line), $R_{F(G)}$ (blue line), M (black line), and $|Q_2^0|$ (green line) on the parameter D for $I/J = 1.2$ and $D/J = 6$.

to as the 1/3 plateau in TLAF); and the saturation region of M in the ferromagnetic phase, which, however, is less pronounced due to anisotropy.

It is also worth noting that as the anisotropy parameter increases, the existence interval of the $\bar{V}$ and V phases in Fig. 8 shrinks. As follows from the phase diagram in Fig. 7, this interval collapses to a point when $D/J \gtrsim 7$.

Fig. 9 presents the dependence of the quantities R_L, $R_{F(G)}$, M and $|Q_2^0|$ on the anisotropy parameter D at a fixed magnetic field $h/J = 1$. These dependencies correspond to movement along the vertical dashed line in Fig. 7.

It can be seen that the quadrupole moment increases as the anisotropy parameter h increases, while the spin moment R_L from the L-sublattice first slightly increases in the region of low fields and then monotonically decreases. The spin moments $R_{F(G)}$ from the F- and G-sublattices do not significantly change throughout the entire range of D. Therefore, the noticeable increase in the total moment M in the ferromagnetic phase is not due to changes in the orientation or absolute values of $R_{F(G)}$, but rather due to the decrease in R_L caused by the anisotropy. Upon transition to the Y-phase, the total moment M begins to decrease, as the rotation of the vectors R_F and R_G around the y-axis reduces their projection onto the z-axis.

The three vertical lines in Fig. 9 divide the four previously described phases. During the transitions from the V-phase to the ferromagnetic phase and from the ferromagnetic phase to the Y-phase, all order parameter dependencies exhibit a kink. At the same time, the transition from the V-phase to the $\bar{V}$-phase is not accompanied by any anomalies in the presented dependencies.

9. GROUND-STATE DEGENERACY AT $I = J$

The case $I = J$ is special because the classical analog of the SU3F Hamiltonian, as we will now show, exhibits continuous accidental degeneracy.

Indeed, let us define the Hamiltonian dependent on the parameter λ:

$$H_\lambda = J\sum_{\{fg\}} S_f S_g + \lambda J\sum_{\{fl\}} S_f S_l + \lambda J\sum_{\{gl\}} S_g S_l + \\ + D\sum_l \left(S_l^y\right)^2 - h\left(\sum_f S_f + \sum_g S_g + \lambda\sum_l S_l\right), \quad (45)$$

where the direction of the magnetic field $h = g\propto_B H$ is generally arbitrary. All the notations in Equation (45) are the same as in the Hamiltonian (1). It is evident that if the conditions $\lambda = I/J = g_L/g$ are met for λ and the field h is directed along the z-axis, the Hamiltonian (45) coincides with the operator H defined by Equation (1).

On the other hand, it is easy to verify that the Hamiltonian (45), up to the constant

$$-JN\left(\frac{3}{2}\lambda^2 S_L(S_L+1) + \frac{9}{4} + \frac{h^2}{6J^2}\right), \quad S_L = 1, \quad (46)$$

can be represented as

$$H_\lambda = D\sum_l \left(S_l^y\right)^2 + \\ + \frac{J}{4}\sum_p \left(S_{pF} + S_{pG} + \lambda S_{pL} - \frac{h}{3J}\right)^2, \quad (47)$$

where the sum p runs over all triangular plaquettes, and the lower indices F, G and L of the spin operators indicate their belonging to the corresponding sublattices in the p-th plaquette.

Thus, if the SU3F parameters satisfy the condition

$$\frac{I}{J} = \frac{g_L}{g}, \quad (48)$$

then the SU3F Hamiltonian in Equation (1) can be represented in the form of Equation (47) with the field h directed along the z-axis.

If we now consider classical moments instead of spin operators in Equation (47), i.e., ordinary vectors of fixed length, it is easy to see that the minimum value of the Hamiltonian (47) will be achieved when both of its terms vanish. The vanishing of the first term implies that the spins of the L-sublattice lie in the easy-plane zx. The requirement for the second term in (47) to vanish reduces to the equation

$$S_{pF} + S_{pG} + \lambda S_{pL} - \frac{h}{3J} = 0. \qquad (49)$$

It is evident that, for certain values of the magnetic fields h, this equation can be satisfied by an infinite set of solutions, i.e., different orientations of the three vectors R_L, R_F and R_G, even when the field h does not lie in the zx-plane. Moreover, if the magnetic field is parallel to the zx-plane (as in our case), the orientation of the vectors R_L, R_F and R_G, which minimizes the Hamiltonian (47), may not necessarily be coplanar with the zx-plane.

The above analysis of the classical limit of the Hamiltonian (47) suggests that the observed (continuous) degeneracy of the SU3F ground state should also hold in the quantum case when the condition (48) is satisfied. Our calculations using the mean-field approximation at $I = J$ and $g_L = g$ confirmed that this is indeed the case.

Similar degeneracy occurs in other quantum magnets, such as the antiferromagnet on a triangular lattice (AFTL) with $S = 1/2$ [43]. As was first demonstrated in [38], this degeneracy can be lifted by considering zero-point quantum fluctuations. This approach requires taking into account higher-order terms (compared to the harmonic approximation used in this work) when bosonizing spin operators within the Holstein–Primakoff representation for the F- and G-subsystems and within the indefinite metric formalism for the L-subsystem.

For this reason, constructing the phase diagram of SU3F at critical parameters satisfying the condition (48) will be carried out by the authors in a separate study.

10. SPIN-WAVE EXCITATIONS IN SU3F UNDER A MAGNETIC FIELD

The spectral properties of SU3F in the absence of a magnetic field were thoroughly studied in [31]. In this section, we analyze changes in the spectrum under a nonzero magnetic field while keeping the anisotropy parameter OA fixed. Four dispersion curves ε_{jk} $(j = 1,\ldots,4)$ were calculated for each set of model parameters based on equation (33) derived in Section 6.

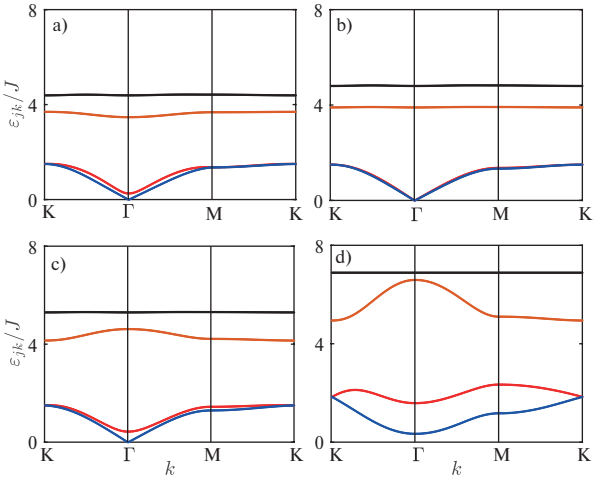

Fig. 10. Spin-wave excitation spectra at $I/J = 0.8$, $D/J = 3$, and four external magnetic field strengths: $h/J = 1$ (*a*), 1.87 (*b*), 3 (*c*), 5.5 (*d*). The wave vector k traverses the triangular path ΓKM in the Brillouin zone (see Fig. 3).

Fig. 10 shows the results of numerical calculations of the dispersion curves for four different magnetic field strengths with the model parameters $I/J = 0.8$ and $D/J = 3$. On the phase diagram in Fig. 4, the four black dots along the horizontal dashed line correspond to these four field values. It is evident that at $h/J = 1$, the system is in the $\bar{Y}$ phase; at $h/J = 1.87$, the system is in the antiparallel phase for the F- and G-sublattices; at $h/J = 3$ the system transitions into the W phase; and at $h/J = 5.5$, the system reaches the ferromagnetic phase. Each of the four panels in Fig. 10 displays four dispersion curves corresponding to the four types of introduced bosons. However, only one curve (black in all graphs) can be confidently associated with the high-energy d-bosons. The other three branches are formed through the hybridization of the a-, b- and c-boson states.

Crucial observation is that in the first three graphs (a, b, c), there is at least one Goldstone mode (blue curves) associated with the breaking of symmetry in the ground state due to the collective rotation of spins in the F- and G-sublattices around the magnetic field direction. In the ferromagnetic phase (Fig. 10d), the ground state does not break this symmetry, and thus, the Goldstone (gapless) mode is absent.

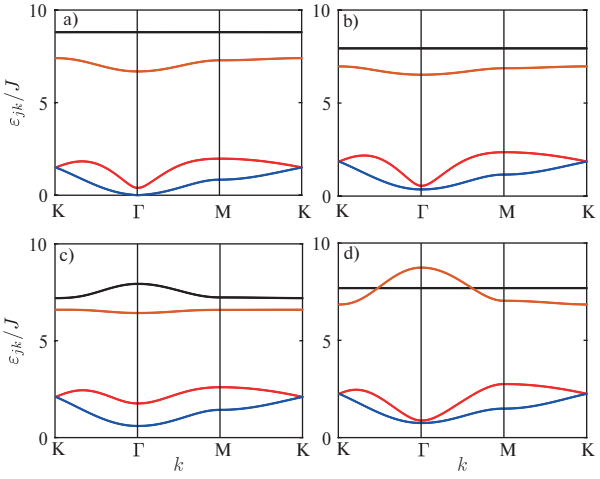

Fig. 11. Spin-wave excitation spectra at $I/J = 1.2$, $D/J = 6$ and four external magnetic field strengths: $h/J = 0.3$ (*a*), 1 (*b*), 4 (*c*), 6 (*d*). The wave vector k traverses the triangular path ΓKM in the Brillouin zone (see Fig. 3).

In Fig. 10b, two Goldstone modes appear (coincident blue and red curves). The origin of the second mode relates to the phase diagram feature discussed in Section 7, specifically the dashed curve (see Fig. 4). In this scenario, the moments R_F and R_G align along the zx -plane in opposite directions, causing the system's energy to degenerate with respect to the rotation of the R_F and R_G vector line around the y-axis.

As mentioned in Section 7, this behavior is due to the vanishing of effective fields and the effective decoupling of the L-subsystem from the F- and G-subsystems. In such a case, the nodes of the L-sublattice become effectively isolated (even from each other), which explains the flat dispersion of the two high-energy branches (black and brown) in Fig. 10b.

The dispersion dependencies ε_{jk} shown in Fig. 11 were calculated with the following model parameters: $I/J = 1.2$, $D/J = 6$, for four values of the external magnetic field: $h/J = 0.3$, 1, 4 and 6. On the phase diagram in Fig. 7, these four field values correspond to the four black dots along the horizontal dashed line. As the magnetic field h increases, the system sequentially transitions through the following four phases: the $\bar{Y}$ phase at $h/J = 0.3$; the ferrimagnetic phase at $h/J = 1$; the V- phase at $h/J = 4.3$; and the ferromagnetic phase at $h/J = 6$.

From the graphs presented in Fig. 11, it is evident that the Goldstone mode appears only in the first case (Fig. 10a), as the breaking of ground-state symmetry (relative to rotations around the z-axis) occurs exclusively in the $\bar{Y}$-phase. In all other regions of the phase diagram (Fig. 7), the spin-wave excitation spectrum remains gapped.

11. CONCLUSION

The main outcome of this study is the construction of the phase diagram of the SU3F ferrimagnet on a triangular lattice, plotted in the coordinates of the magnetic field h (applied in the easy-plane anisotropy plane) and the single-ion anisotropy parameter D at zero temperature. Among the key features of the SU3F model, the following three stand out: 1) different spin magnitudes in magnetic sublattices: two sublattices (F and G) have spin $S = 1/2$, while the third L-sublattice has spin $S = 1$; 2) single-ion anisotropy: easy-plane anisotropy acts on the L-sublattice with spin S=1; 3) different exchange integrals: The exchange interactions differ between the F- and G-sublattices (J) and between the L- F(G) sublattices (I).

Numerical calculations under the mean-field approximation revealed two qualitatively distinct types of SU3F phase diagrams depending on the ratio between the exchange integrals I and J. These phase diagrams differ both in the number of realized phases and in the nature of their magnetic structures.

For $I < J$, the ground state of SU3F can be characterized by three magnetic configurations: the $\bar{Y}$, W phase, and the ferromagnetic phase (see Fig. 4). Notably, along the boundary between the $\bar{Y}$ - and W -phases (dashed line in Fig. 4), the SU3F system effectively splits into two independent magnetic subsystems/ The first subsystem consists of spin-1 sites on a triangular lattice and behaves as a paramagnet. Another one consists of $S = 1/2$ spins forming a planar hexagonal lattice in a collinear two-sublattice antiferromagnetic phase under an effective zero magnetic field. This decoupling leads to an additional degeneracy of the ground state, associated with the free rotation of the antiferromagnetic vector within the easy-plane. This degeneracy manifests as an extra Goldstone mode in the spin-wave excitation spectrum.

For the reverse exchange ratio ($I > J$), the SU3F the h–D- phase diagram undergoes significant changes. It now features four distinct regions characterized by different magnetic ground-state structures, i.e. the Y- phase, two collinear ferrimagnetic and ferromagnetic phases, as well as the V-phase. The V-phase can further be subdivided into two sub-phases ($\bar{V}$ and V), depending on whether the angle θ_L exceeds the critical value $\pi/2$.

For both $I < J$ and $I > J$, the quadrupole and dipole order parameters were analyzed as functions of both the magnetic field (at fixed anisotropy OA) and the anisotropy parameter OA (at fixed magnetic field h). One significant result of this study is the dependence of the total moment M on the external magnetic field. For $I > J$ and a specific finite value of OA, this dependence qualitatively reproduces the well-known behavior observed in quantum antiferromagnets on a triangular lattice with uniform spin $S = 1/2$ and without anisotropy [38, 43]. Specifically, within a certain magnetic field range, the magnetization curve exhibits a plateau (albeit with a slight tilt in our case). In conventional quantum triangular-lattice antiferromagnets (QTAFMs), this plateau arises due to quantum antiferromagnetic fluctuations, while in SU3F, it emerges due to the presence of single-ion anisotropy.

A notable finding is the qualitative difference between the two phase diagrams for $I < J$ and $I > J$. There is no continuous transformation at $I \to J$ from one diagram to the other. This is because, when the exchange integrals become equal ($I = J$), accidental degeneracy arises, leading to an ambiguity in the magnetic configuration within the mean-field approximation for given magnetic field and anisotropy values. We hypothesize that, as with QTAFMs, quantum fluctuations should lift the observed accidental degeneracy (as well as the additional degeneracy noted for $I < J$). However, a detailed investigation of this issue requires further study and will be addressed in future research.

In conclusion, we emphasize that in the present study, the magnetic field h applied to the quantum SU3F system was oriented within the easy-plane anisotropy plane. If the magnetic field were instead applied perpendicular to this plane, the behavior of the magnetic order parameters could differ qualitatively.

ACKNOWLEDGMENTS

The authors express their gratitude to Professor V.V. Valkov for his assistance in formulating the research problem, providing valuable advice, and stimulating discussions regarding the obtained results.

FUNDING

This study was conducted within the framework of the scientific program of the State Assignment of the Institute of Physics of the Siberian Branch of the Russian Academy of Sciences (IF SB RAS).

APPENDIX A. UNITARY TRANSFORMATION OF HUBBARD OPERATORS

As a result of the unitary transformations of the Hubbard operators according to formula (20), with the unitary operator $U_{\tilde{n}\tilde{m}}(\alpha)$ ($n \neq m$)defined by formula (20), the following expressions are obtained [40]:

$$X^{nn} = \cos^2\alpha X^{\tilde{n}\tilde{n}} + \sin^2\alpha X^{\tilde{m}\tilde{m}} - \\ -\frac{1}{2}\sin 2\alpha\left(X^{\tilde{n}\tilde{m}} + X^{\tilde{m}\tilde{n}}\right),$$

$$X^{mm} = \cos^2\alpha X^{\tilde{m}\tilde{m}} + \sin^2\alpha X^{\tilde{n}\tilde{n}} + \\ +\frac{1}{2}\sin 2\alpha\left(X^{\tilde{n}\tilde{m}} + X^{\tilde{m}\tilde{n}}\right),$$

$$X^{nm} = \cos^2\alpha X^{\tilde{n}\tilde{m}} - \sin^2\alpha X^{\tilde{m}\tilde{n}} + \\ +\frac{1}{2}\sin 2\alpha\left(X^{\tilde{n}\tilde{n}} - X^{\tilde{m}\tilde{m}}\right),$$

$$X^{mn} = \cos^2\alpha X^{\tilde{m}\tilde{n}} - \sin^2\alpha X^{\tilde{n}\tilde{m}} + \\ +\frac{1}{2}\sin 2\alpha\left(X^{\tilde{n}\tilde{n}} - X^{\tilde{m}\tilde{m}}\right),$$

$$X^{np} = \cos\alpha X^{\tilde{n}\tilde{p}} - \sin\alpha X^{\tilde{m}\tilde{p}},$$

$$X^{pn} = \cos\alpha X^{\tilde{p}\tilde{n}} - \sin\alpha X^{\tilde{p}\tilde{m}},$$

$$X^{pm} = \cos\alpha X^{\tilde{p}\tilde{m}} + \sin\alpha X^{\tilde{p}\tilde{n}},$$

$$X^{mp} = \cos\alpha X^{\tilde{m}\tilde{p}} + \sin\alpha X^{\tilde{n}\tilde{p}},$$

$$X^{pq} = X^{\tilde{p}\tilde{q}},$$

where all four state indices p, q, n and m are different, and the site indices are omitted. In the main text, for brevity, the tilde notation, indicating the new (transformed) states, is not used for the indices of the thrice-transformed Hubbard operators.

APPENDIX B. MATRIX ELEMENTS OF SPIN OPERATORS

This appendix presents the explicit form of the matrix elements $s^{\alpha}_{nm} \equiv \langle n \mid S^{\alpha}_{l} \mid m \rangle$ ($\alpha = \{x,y,z\}$ and $n,m = \{\bar{1},0,1\}$),, used in the decomposition (27). These elements were obtained from the three successive transformations of the Hubbard operators using the three unitary operators $U_{1-1}(-\alpha_1)$, $U_{0-1}(-\alpha_3)$ and $U_{10}(-\alpha_2)$, followed by substituting

the transformed results into the representation (15) for the spin operators of the L -sublattice.

Matrix elements for the spin operator S_l^z :

$$s_{11}^z = (\cos\alpha_1\cos\alpha_2 + \sin\alpha_1\sin\alpha_2\sin\alpha_3)^2 -$$

$$-\sin^2\alpha_1\cos^2\alpha_3,$$

$$s_{\bar{1}\bar{1}}^z = (\cos\alpha_1\sin\alpha_2\sin\alpha_3 - \sin\alpha_1\cos\alpha_2)^2 -$$

$$-\cos^2\alpha_1\cos^2\alpha_3,$$

$$s_{00}^z = \sin^2\alpha_2\cos^2\alpha_3 - \sin^2\alpha_3,$$

$$s_{10}^z = s_{01}^z = -\frac{1}{2}\sin\alpha_1(1+\sin^2\alpha_2)\sin 2\alpha_3 -$$

$$-\frac{1}{2}\cos\alpha_1\sin(2\alpha_2)\cos\alpha_3,$$

$$s_{\bar{1}0}^z = s_{0\bar{1}}^z = -\frac{1}{2}\cos\alpha_1(1+\sin^2\alpha_2)\sin 2\alpha_3 +$$

$$+\frac{1}{2}\sin\alpha_1\sin(2\alpha_2)\cos\alpha_3,$$

$$s_{\bar{1}1}^z = s_{1\bar{1}}^z = \frac{1}{2}\cos 2\alpha_1\sin 2\alpha_2\sin\alpha_3 +$$

$$+\frac{1}{2}\sin 2\alpha_1(\sin^2\alpha_2\sin^2\alpha_3 - \cos^2\alpha_3 - \cos^2\alpha_2).$$

For the operator S_l^x :

$$s_{11}^x = \sqrt{2}(\cos\alpha_1\sin\alpha_2 -$$

$$-\sin\alpha_1\sin\alpha_3\cos\alpha_2)(\cos\alpha_1\cos\alpha_2 +$$

$$+\sin\alpha_1\sin\alpha_3\sin\alpha_2 + \sin\alpha_1\cos\alpha_3),$$

$$s_{\bar{1}\bar{1}}^x = \sqrt{2}(\sin\alpha_1\sin\alpha_2 +$$

$$+\cos\alpha_1\sin\alpha_3\cos\alpha_2)(\sin\alpha_1\cos\alpha_2 -$$

$$-\cos\alpha_1\sin\alpha_3\sin\alpha_2 - \cos\alpha_1\cos\alpha_3),$$

$$s_{00}^x = \frac{1}{\sqrt{2}}(\cos\alpha_2\sin 2\alpha_3 - \sin 2\alpha_2\cos^2\alpha_3),$$

$$s_{\bar{1}1}^x = \frac{\cos 2\alpha_1}{\sqrt{2}}(\sin\alpha_2\cos\alpha_3 - \sin\alpha_3\cos 2\alpha_2) -$$

$$-\frac{\sin 2\alpha_1}{2\sqrt{2}}(\cos\alpha_2\sin 2\alpha_3 + \sin 2\alpha_2(1+\sin^2\alpha_3)),$$

$$s_{10}^x = \frac{\cos\alpha_1}{\sqrt{2}}(\cos 2\alpha_2\cos\alpha_3 + \sin\alpha_2\sin\alpha_3) +$$

$$+\frac{\sin\alpha_1}{\sqrt{2}}(\cos\alpha_2\cos 2\alpha_3 + \frac{1}{2}\sin 2\alpha_2\sin 2\alpha_3),$$

$$s_{\bar{1}0}^x = -\frac{\sin\alpha_1}{\sqrt{2}}(\cos 2\alpha_2\cos\alpha_3 + \sin\alpha_2\sin\alpha_3) +$$

$$+\frac{\cos\alpha_1}{\sqrt{2}}(\cos\alpha_2\cos 2\alpha_3 + \frac{1}{2}\sin 2\alpha_2\sin 2\alpha_3),$$

$$s_{\bar{1}1}^x = s_{1\bar{1}}^x, \quad s_{10}^x = s_{01}^x, \quad s_{\bar{1}0}^x = s_{0\bar{1}}^x.$$

For the operator S_l^y :

$$s_{11}^y = s_{\bar{1}\bar{1}}^y = s_{00}^y = 0,$$

$$s_{01}^y = \frac{i}{\sqrt{2}}(-\sin\alpha_1\cos\alpha_2 +$$

$$+\cos\alpha_1(\cos\alpha_3 + \sin\alpha_2\sin\alpha_3)),$$

$$s_{\bar{1}0}^y = \frac{i}{\sqrt{2}}(\cos\alpha_1\cos\alpha_2 +$$

$$+\sin\alpha_1(\cos\alpha_3 + \sin\alpha_2\sin\alpha_3)),$$

$$s_{\bar{1}1}^y = \frac{i}{\sqrt{2}}(\sin\alpha_2\cos\alpha_3 - \sin\alpha_3),$$

$$s_{\bar{1}1}^y = -s_{1\bar{1}}^y, \quad s_{10}^y = -s_{01}^y, \quad s_{\bar{1}0}^y = -s_{0\bar{1}}^y.$$

For the operator $(S_l^y)^2$:

$$\langle 1 \mid (S_l^y)^2 \mid 1\rangle = \frac{1}{2} + \frac{1}{2}\cos^2\alpha_1\sin^2\alpha_2 +$$

$$+\frac{1}{2}\sin^2\alpha_1(\cos^2\alpha_2\sin^2\alpha_3 - \sin\alpha_2\sin 2\alpha_3) -$$

$$-\frac{1}{2}\sin 2\alpha_1\cos\alpha_2(\sin\alpha_2\sin\alpha_3 + \cos\alpha_3),$$

$$\langle \bar{1} \mid (S_l^y)^2 \mid \bar{1}\rangle = \frac{1}{2} + \frac{1}{2}\sin^2\alpha_1\sin^2\alpha_2 +$$

$$+\frac{1}{2}\cos^2\alpha_1(\cos^2\alpha_2\sin^2\alpha_3 - \sin\alpha_2\sin 2\alpha_3) +$$

$$+\frac{1}{2}\sin 2\alpha_1\cos\alpha_2(\sin\alpha_2\sin\alpha_3 + \cos\alpha_3),$$

$$\langle 0 \mid (S_l^y)^2 \mid 0\rangle = \frac{1}{2}(\sin\alpha_2\sin 2\alpha_3 +$$

$$+1 + \cos^2\alpha_2\cos^2\alpha_3),$$

$$\langle \bar{1} \mid (S_l^y)^2 \mid 1\rangle = \frac{1}{4}(\cos^2\alpha_2\sin^2\alpha_3 -$$

$$-\sin\alpha_2\sin 2\alpha_3 - \sin^2\alpha_2)\sin 2\alpha_1 -$$

$$-\frac{1}{2}\cos 2\alpha_1\cos\alpha_2(\sin\alpha_2\sin\alpha_3 + \cos\alpha_3),$$

$$\langle \bar{1} \mid (S_l^y)^2 \mid 0\rangle = \frac{1}{2}\sin\alpha_1\cos\alpha_2(\sin\alpha_3 - \sin\alpha_2\cos\alpha_3) +$$

$$+\frac{1}{2}\cos\alpha_1\left(\sin\alpha_2\cos 2\alpha_3 - \frac{1}{2}\cos^2\alpha_2\sin(2\alpha_3)\right),$$

$$\langle 1 \mid (S_l^y)^2 \mid 0\rangle = \frac{1}{2}\cos\alpha_1\cos\alpha_2(\sin\alpha_2\cos\alpha_3 - \sin\alpha_3) +$$

$$+\frac{1}{2}\sin\alpha_1\left(\sin\alpha_2\cos 2\alpha_3-\frac{1}{2}\cos^2\alpha_2\sin(2\alpha_3)\right),$$

$$\langle 1|(S_l^y)^2|\bar{1}\rangle=\langle\bar{1}|(S_l^y)^2|1\rangle,\quad \langle 0|(S_l^y)^2|\bar{1}\rangle=\langle\bar{1}|(S_l^y)^2|0\rangle,$$

$$\langle 0\mid(S_l^y)^2\mid 1\rangle=\langle 1\mid(S_l^y)^2\mid 0\rangle.$$

APPENDIX C. BOSONIZATION OF SPIN OPERATORS FOR S=1

Using the representation (28) in formulas (27) and retaining only terms up to the second order in boson operators, the following spin operator expressions through bosonic operators are obtained:

$$S_l^x=\frac{1}{\sqrt{2}}[(s_{0,1}^x(c_l^+ + c_l)+s_{\bar{1},1}^x(d_l^+ + d_l)+$$

$$+s_{\bar{1},0}^x(d_l^+c_l + c_l^+d_l)+s_{\bar{1},1}^x+(s_{0,0}^x - s_{1,1}^x)c_l^+c_l +$$

$$+(s_{\bar{1},\bar{1}}^x - s_{1,1}^x)d_l^+d_l],$$

$$S_l^y=\frac{i}{\sqrt{2}}[s_{0,1}^y(c_l^+ - c_l)+s_{\bar{1},1}^y(d_l^+ - d_l)+$$

$$+s_{\bar{1},0}^y(d_l^+c_l - c_l^+d_l)],$$

$$S_l^z=s_{0,1}^z(c_l^+ + c_l)+s_{\bar{1},1}^z(d_l^+ + d_l)+$$

$$+s_{\bar{1},0}^z(d_l^+c_l + c_l^+d_l)+s_{\bar{1},1}^z+(s_{0,0}^z - s_{1,1}^z)c_l^+c_l +$$

$$+(s_{\bar{1},\bar{1}}^z - s_{1,1}^z)d_l^+d_l,$$

$$(S_l^y)^2=\frac{1}{2}[((s_{\bar{1},0}^y)^2-(s_{\bar{1},1}^y)^2)c_l^+c_l +$$

$$+((s_{\bar{1},0}^y)^2-(s_{0,1}^y)^2)d_l^+d_l -$$

$$-s_{\bar{1},0}^y s_{0,1}^y(d_l^+ + d_l)+s_{\bar{1},0}^y s_{\bar{1},1}^y(c_l^+ + c_l)+$$

$$+((s_{0,1}^y)^2+(s_{\bar{1},1}^y)^2)+s_{0,1}^y s_{\bar{1},1}^y(d_l^+c_l + c_l^+d_l)].$$

The presented expressions, after averaging and applying the spectral theorem to compute the boson operator expectations, were used to derive the formulas for calculating the order parameters R_L, M and Q_2^0.

REFERENCES

1. *B. Barbara, Y. Imry, G. Sawatzky, and P. C. E. Stamp*, Quantum Magnetism NATO: Science for Peace and Security, Series B: Physics and Biophysics, Springer Dordrecht (2008).
2. *Introduction to Frustrated Magnetism*, Springer Series in Solid-State Sciences, ed. by C. Lacroix, P. Mendels, and F. Mila (2011).
3. *A. Auerbach*, Interacting Electrons and Quantum Magnetism, Springer-Verlag, New York, Inc. (1994).
4. *V. V. Val'kov, S. G. Ovchinnikov*, Theoretical and Mathematical Physics, **50**, 306 (1982).
5. *H. H. Chen and P. M. Levy*, Phys. Rev. Lett. **27**, 1383 (1971).
6. *V. M. Matveev*, JETP **38**, 813 (1973).
7. *M. P. Kashchenko, N. F. Balakhonov, and L. V. Kurbatov*, JETP **37**, 201 (1973).
8. *V. M. Loktev and V. S. Ostrovskii*, Ukr. J. Phys. **23**, 1717 (1978).
9. *F. P. Onufrieva*, JETP **53**, 1241 (1981).
10. *A. F. Andreev and I. A. Grishchuk*, JETP **60**, 267 (1984).
11. *F. P. Onufrieva*, JETP **62**, 1311 (1985).
12. *N. Papanicolaou*, Nucl. Phys. B **305**, 367 (1988).
13. *A. V. Chubukov*, J. Phys.: Condens. Matter **2**, 1593 (1990).
14. *V. V. Val'kov and T. A. Val'kova*, JETP **72**, 1053 (1991).
15. *Yu. A. Fridman, O. A. Kosmachev, and Ph. N. Klevets*, J. Magn. Magn. Mat. **325**, 125 (2013).
16. A. I. Smirnov, Physics Uspekhi 59, 564 (2016)].
17. *O. A. Kosmachev, Ya. Yu. Matyunina, and Yu. A. Fridman*, JETP **135**, 354 (2022).
18. *H. F. Verona de Resende, F. C. SaBarreto, and J. A. Plascak*, Physica A **149**, 606 (1988).
19. *G. M. Zhang and C. Z. Yang*, Phys. Rev. B **48**, 9452 (1993).
20. *A. Bobak and M. Jurcisin*, Physica A *240*, 647 (1997)
21. *G. M. Buendia and M. A. Novotny*, J. Phys.: Condens. Matter **9**, 5951 (1997).
22. *M. Godoy and W. Figueiredo*, Physica A **339**, 392 (2004).
23. *T. Iwashita and N. Uryu*, J. Phys. Soc. Japan **53**, 721 (1984).
24. *J. Oitmaa*, Phys. Rev. B **72**, 224404 (2005).
25. *J. Oitmaa and I. G. Enting*, J. Phys.: Condens. Matter **18**, 10931 (2006).
26. *W. Selke and J. Oitmaa*, J. Phys.: Condens. Matter **22**, 076004 (2010).
27. *Yu. A. Fridman and O. A. Kosmachev*, Physics of the Solid State **51**, 1167 (2009).
28. *M. Zukovic and A. Bobak*, Phys. Rev. E **91**, 052138 (2015).
29. *M. Zukovic and A. Bobak*, Physica A **436** 509 (2015).
30. *E. S. de Santana, A. S. de Arruda, and M. Godoy*, Condensed Matter Physics, **26**, 23601 (2023)
31. *A. S. Martynov and D. M. Dzebisashvili*, J. Magn. Magn. Mat. **584**, 171906 (2024).

32. *B. A. Ivanov and A. K. Kolezhuk*, Phys. Rev. B **68**, 052401 (2003).

33. *A. Lauchli, F. Mila, and K. Penc*, Phys. Rev. Lett. **97**, 087205 (2006).

34. *V. V. Val'kov and M. S. Shustin*, JETP **121**, 860 (2015)].

35. *V. V. Val'kov and M. S. Shustin*, J. Low Temp. Phys. **185**, 564 (2016).

36. *Yu. A. Fridman and D. V. Spirin*, J. Magn. Magn. Mat. **253**, 111 (2002).

37. *J. Hubbard*, Proc. Roy. Soc. A **281**, 401 (1964).

38. *A. V. Chubukov and D. I. Golosov*, J. Phys.: Condens. Matter ***3***, 69 (1991).

39. *C. Cohen-Tannoudji, B. Diu, and F. Laloe*, Wiley-VCH (2019), Vol. 1.

40. *V. V. Val'kov*, Theor. Math. Phys. **76**, 766 (1988).

41. *A. I. Akhiezer, V. G. Bar'yakhtar, and S. V. Peletminskii*, Spin Waves, Published by North-Holland Publishing Company, Amsterdam, 1968.

42. *L. D. Landau and E. M. Lifshitz*, Quantum Mechanics: Non-Relativistic Theory, Elsevier (2013), Vol. 3.

43. *T. Coletta, T. A. Toth, K. Penc at al.*, Phys. Rev. B **94**, 075136 (2016).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 110–123

ORDER, DISORDER, AND PHASE TRANSITION IN CONDENSED MEDIA

STATIONARY AND NON STATIONARY CURRENT IN FINITE KITAEV CHAINS

© 2025 Yu. M. Bilinskii[a], P. I. Arseev[a,*], N. S. Maslova[b]

[a]Lebedev Physical Institute, Russian Academy of Sciences, Moscow, Russia
[b]Lomonosov Moscow State University, Moscow, Russia
*e-mail: ars@lpi.ru

Received November 08, 2024

Revised November 18, 2024

Accepted November 18, 2024

Abstract. We investigate the role of gap states in processes of charge transmission along finite superconducting Kitaev chain. We use the formalism of non-stationary Green's functions, which contain full information about the non-equilibrium and non-stationary properties of the system. We discuss tunneling current and non-stationary transport properties of a finite Kitaev chain in the subgap regime. Under the assumption that the finite Kitaev chain is connected at each edge to its own external lead (normal reservoir) we obtain time-dependent behavior of the tunneling current after the sudden change of bias voltage in one of the leads. Obtained results show how quickly the "Majorana mode" at one edge of the chain responds after external perturbation acts on the "Majorana mode" at the other edge. Presented calculations are completely analytical and straightforward, in contrast with many other methods.

Keywords: *Kitaev superconducting chain, tunneling in superconductors*

DOI: 10.31857/S00444510250110e7

1. INTRODUCTION

In recent decades, significant attention has been given to systems exhibiting "topologically nontrivial" properties. However, for practical applications, it is essential to assess the specific physical characteristics of such systems in addition to their mathematical interpretation of the ground state properties. One of the simplest models that demonstrates properties allowing for topological interpretation is the atomic chain with ppp-wave superconductivity of spinless particles, proposed by A. Kitaev [1]. The primary interest in this model in subsequent years was driven by the nontrivial topological interpretation of its ground-state properties. It was shown that, due to "topological reasons", quantum states localized at the chain edges appear within the superconducting gap. These states, often referred to as "Majorana modes", are commonly associated with the existence of quasiparticles [2] that bear resemblance to Majorana fermions [3].

Possible experimental realizations of this model are typically based on the proximity effect in semiconductor nanowires with strong spin-orbit interaction, placed on a superconducting substrate [4, 5, 6]. The latest experimental advances and discussions on the challenges encountered can be found in the review [7].

It is widely believed that further progress in this field may involve models with an effective Josephson action, accounting for Coulomb blockade-type effects [8, 9, 10, 11, 12, 13]. There is hope that long-range Coulomb interactions could facilitate signal transmission in finite Kitaev chains using "Majorana states". However, recharging effects inevitably involve charge transfer processes, so we must ensure that we accurately describe tunneling transport and charge transfer effects first in the simplest tunneling setup. Theoretical results can then be compared with tunneling experiments under various conditions [14, 15].

Some theoretical studies suggest that "Majorana states" could be utilized as an error-protected method for storing and transmitting information in quantum technology [16, 17]. However, if a state is protected from arbitrary changes due to external

noise, the same protection may make intentional changes to the system's state equally difficult, potentially rendering the system impractical for real-world applications. One possible way to study how well a system responds to a signal is to investigate its nonstationary transport properties.

In [18], nonstationary effects related to tunnel barrier transparency modulation were considered in a quasiclassical approach. That work analyzed a three-terminal system, where one of the contacts was effectively used to fix the chemical potential of the superconductor. In this study, we consider a two-terminal geometry, where the superconductor is connected only to two external contacts. To explore the role of localized states in nonstationary transport properties, we employ the formalism of nonstationary Green's functions for electrons.

Below, we will demonstrate that this approach enables us to derive explicit analytical expressions for both the tunnel current and nonstationary charge transport, in contrast to more complex methods based on density matrix equations, as discussed, for example, in [19]. Furthermore, this approach allows us to compare quasiclassical calculations with microscopic methods and establish a connection between the parameters used in these different approaches.

The exact electronic Green's functions for the infinite Kitaev chain in equilibrium can be obtained analytically [20]. These functions can be used to derive the nonstationary Green's functions for a finite chain, allowing us to understand how the system evolves over time when subjected to an external perturbation. The key idea in our approach is to treat the finite Kitaev chain as a cut segment of an infinite chain or as a chain with strong defects (for a single-cut chain, see, e.g., [21]). This trick enables us to use the Green's functions of the infinite chain to study all single-particle states in the system. Our calculations do not require any special interpretation of singularities in the single-particle Green's function as specific "states". It is important to note that the poles of the single-particle Green's function, which appear inside the superconducting gap in this model, can hardly be interpreted as single-particle excitations. True Majorana particles, as discussed in the pioneering works [3], are well-defined particles (quasiparticles) with the usual algebra of creation and annihilation operators. In any physical problem, such real particles contribute to the single-particle Green's function with a residue equal to one. It is well known that bound states localized around defects, such as paramagnetic impurities [22] or resonance impurities [23] with energies lying inside the superconducting gap, frequently appear in conventional superconductors. These states are genuine single-particle states. In the present case, we observe that the appearance of poles in the electronic Green's function within the gap, with residues smaller than one, is more likely an artifact of the model, which has a degenerate (in the highly symmetric case) ground state, rather than the emergence of new quasiparticles.

2. PROPERTIES OF AN ISOLATED KITAEV CHAIN

In this section, we briefly reproduce some results related to the spectral properties of a finite Kitaev chain, using the Green's function formalism, which we will employ in later sections.

We start with the free ideal Kitaev chain, which is completely isolated from any external systems.

The model Hamiltonian of such a system can be written as

$$\widehat{\widehat{H}} = -\mu\sum_{n=1}^{N}\psi_n\psi_n - t\sum_{n=1}^{N-1}\left(\psi_n\psi_{n+1} + \psi_{n+1}\psi_n\right) + \\ + \sum_{n=1}^{N-1}\left(\Delta\psi_n\psi_{n+1} + \Delta^*\psi_{n+1}\psi_n\right). \quad (1)$$

Here, ψ_n and ψ_n are the creation and annihilation operators for a particle at site n, μ is the chemical potential, t is the hopping parameter between neighboring sites, Δ is the superconducting order parameter, which in this study we consider as a fixed parameter, N is the total number of sites in the lattice.

To obtain exact solutions for the Green's functions of the Hamiltonian (1), it is convenient to use the Green's functions of an infinite Kitaev chain. Indeed, the behavior of a finite chain can be modeled by considering an infinite chain with infinitely strong point defects $U\to+\infty$ added at sites 0 and N+1 (see Fig. 1). As a result, the particles located between these two sites will be completely isolated from the outer parts of the chain, and the Green's functions will be identical to those of a finite Kitaev chain of length N, as long as the node indices lie between 0 and N+1.

Thus, the behavior of the system is described by the following Hamiltonian, corresponding to the system shown in the figure:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}, \tag{2}$$

where

$$\widehat{H}_0 = -\mu\sum_n \psi_n\psi_n - t\sum_n\left(\psi_n\psi_{n+1} + \psi_{n+1}\psi_n\right) + \\ +\sum_n\left(\Delta\psi_n\psi_{n+1} + \Delta^*\psi_{n+1}\psi_n\right),$$

$$\widehat{V} = U\left(\psi_0\psi_0 + \psi_{N+1}\psi_{N+1}\right).$$

This Hamiltonian (2) is identical to Hamiltonian (1) when $U\to\infty$. To determine the physical properties of the chain, we use the formalism of normal and anomalous Green's functions, denoted as $G_{nm}(t,t')$, $F_{nm}(t,t')$, respectively. In this work, we use the following definitions of Green's functions:

$$\Gamma^R_{nm}(t,\,t') = \begin{pmatrix} G^R_{nm}(t,\,t') & F^R_{nm}(t,\,t') \\ F^{R+}_{nm}(t,\,t') & G^{R+}_{nm}(t,\,t') \end{pmatrix} = -i\left\langle\begin{pmatrix} \left\{\psi_n(t),\,\psi^+_m(t')\right\} & \left\{\psi_n(t),\,\psi_m(t')\right\} \\ \left\{\psi^+_n(t),\,\psi^+_m(t')\right\} & \left\{\psi^+_n(t),\,\psi_m(t')\right\} \end{pmatrix}\right\rangle\theta(t-t'), \tag{3}$$

$$\Gamma^A_{nm}(t,\,t') = \begin{pmatrix} G^A_{nm}(t,\,t') & F^A_{nm}(t,\,t') \\ F^{A+}_{nm}(t,\,t') & G^{A+}_{nm}(t,\,t') \end{pmatrix} = i\left\langle\begin{pmatrix} \left\{\psi_n(t),\,\psi^+_m(t')\right\} & \left\{\psi_n(t),\,\psi_m(t')\right\} \\ \left\{\psi^+_n(t),\,\psi^+_m(t')\right\} & \left\{\psi^+_n(t),\,\psi_m(t')\right\} \end{pmatrix}\right\rangle\theta(t'-t), \tag{4}$$

$$\Gamma^<_{nm}(t,\,t') = \begin{pmatrix} G^<_{nm}(t,\,t') & F^<_{nm}(t,\,t') \\ F^{<+}_{nm}(t,\,t') & G^{<+}_{nm}(t,\,t') \end{pmatrix} = -i\left\langle\begin{pmatrix} \psi^+_m(t')\psi_n(t) & \psi_m(t')\psi_n(t) \\ \psi^+_m(t')\psi^+_n(t) & \psi_m(t')\psi^+_n(t) \end{pmatrix}\right\rangle, \tag{5}$$

where $\left\{\hat{a},\hat{b}\right\} = \hat{a}\hat{b} + \hat{b}\hat{a}$, $<\hat{a}>= \mathrm{Tr}(\hat{\rho}\hat{a})$. The indices R and A denote the retarded and advanced Green's functions, respectively.

Using Dyson's equation for Hamiltonian (2), we can express the retarded Green's functions $\Gamma^R_{nm}(t,t')$ of the finite chain in terms of the Green's functions $\Gamma^{0R}_{nm}(t,t')$ of the infinite Kitaev chain. In the Appendix, it is shown that the functions $\Gamma^R_{nm}(\omega)$ have poles at points $\omega = \pm\omega_0$, where

$$\omega_0 = \frac{|\Delta|(4t^2-\mu^2)}{it\sqrt{4(t^2-|\Delta|^2)-\mu^2}}\left(\chi_+^{N+1} - \chi_-^{N+1}\right). \tag{6}$$

Here

$$\chi_\pm = \frac{-\mu \pm i\sqrt{4t^2 - \left(\mu^2 + 4\,|\,\Delta\,|^2\right)}}{2(t+|\Delta|)}. \tag{7}$$

As $|\,\chi_\pm\,|<1$ (see 52), expression (6) is written for the case $|\chi_\pm|^N \ll 1$. For sufficiently large N, the parameter ω_0 is small compared to other system parameters and decays exponentially as the N chain length increases.

For the cases $|\Delta| \ll t$ and $|\Delta| < t$, $|\Delta| \to t$, the "exponential smallness" of expression (6) in N can be explicitly demonstrated:

$$\omega_0 = \begin{cases} 4|\Delta|e^{-N(|\Delta|/t)}, & |\Delta| << t, \\ 2te^{-N\ln\left(\sqrt{2t/(t-|\Delta|)}\right)}, & (t-|\Delta|) \ll t. \end{cases} \tag{8}$$

The exponential decay of ω_0 with increasing chain length is explained by the exponentially weak overlap of the two bound states at opposite edges of the chain. Using (8), we can estimate the localization length of the bound states as:

$$l_{loc} \simeq \begin{cases} a(t/|\Delta|), & |\Delta| << t, \\ a/\ln\left(\sqrt{2t/(t-|\Delta|)}\right), & (t-|\Delta|) \ll t, \end{cases} \tag{9}$$

where a is the lattice constant.

Such an exponential dependence has been observed in tunneling experiments using Coulomb blockade methods, as described in [15].

In the limit $N\to\infty$, the states near each edge begin to behave as if the chain were semi-infinite. In this case, the two poles with residues equal to 1/2 together correspond to a single Fermi excitation, which is split between the two edges of the chain. Thus, the residue in terms of Bogoliubov excitations is equal to 1, as it should be. However, when observing only one end of the chain, we "see" only half of this excitation. This Fermi excitation is very specific because it is

Fig. 1: Infinite Kitaev chain with two defects

the excitation that connects two degenerate ground states with different parity (i.e., a different number of electrons), but with the same energy.

This statement can be easily illustrated with a simple example of a two-site chain. The Hamiltonian (1) for two sites can be diagonalized using the Bogoliubov transformation. In terms of Bogoliubov operators, the Hamiltonian takes the form

$$\widehat{H}_2 = E_0 + \varepsilon_1 c_1^+ c_1 + \varepsilon_2 c_2^+ c_2, \qquad (10)$$

where

$$c_{1,2} = u\frac{(\psi_1 + \psi_2)}{\sqrt{2}} \pm v\frac{(\psi_1^+ - \psi_2^+)}{\sqrt{2}},$$

$$v^2, u^2 = \frac{1}{2}\left(1 \pm \frac{\mu}{t}\right),$$

$$\varepsilon_{1,2} = t\left(1 \mp \frac{\Delta}{\sqrt{t^2 - \mu^2}}\right).$$

For $\Delta = \sqrt{t^2 - \mu^2}$ (which in the case $\mu = 0$ gives $\Delta = t$), we obtain $\varepsilon_1 = 0$ (the solution 53 for the case $N = 2$, $\omega = 0$). Then, $|\Phi_0> = \frac{1}{\sqrt{2}}(\psi_1^+ - \psi_2^+)|0>$ corresponds to the ground state and satisfies $c_{1,2}|\Phi_0> = 0$. At the same time, the state $|\Phi_1> = c_1|\Phi_0> = (v + u\psi_1^+\psi_2^+)|0>$ also has zero energy, which means that the ground state is degenerate. For the matrix elements between these ground states, we have

$$<\Phi_0|\psi_1|\Phi_1> = u/\sqrt{2}, <\Phi_1|\psi_1|\Phi_0> = v/\sqrt{2}.$$

This means that in the single-particle function G_{11} at $\omega = 0$, a pole appears with a residue equal to 1/2.

3. TUNNELING CURRENT

We first consider the stationary tunneling properties of the Kitaev chain. To do this, we assume that the chain is connected at sites 1 and N to two external reservoirs with a large number of degrees of freedom, labeled by indices l and r, respectively.

The total Hamiltonian can then be written as

$$\begin{aligned}\widehat{\widehat{H}} = \ & \widehat{H} + \sum_p \tau_p^l \left(h_p^{l+}\psi_1 + \psi_1^+ h_p^l\right) + \\ & + \sum_p \tau_p^r \left(h_p^{r+}\psi_N + \psi_N^+ h_p^r\right) + \\ & + \sum_p E_p^l h_p^{l+} h_p^l + \sum_p E_p^r h_p^{r+} h_p^r.\end{aligned} \qquad (11)$$

The current flowing into the chain through site 1 is given by the standard expression ([24]):

$$I_l(t) = i\sum_p \tau_p^l < h_p^{l+}\psi_1 - \psi_1^+ h_p^l > . \qquad (12)$$

Using the nonstationary diagrammatic technique, this expression can be rewritten as

$$I_l(t) = -\sum_p \tau_p^l \left(\tilde{G}^<_{lp,1}(t,t) - \tilde{G}^<_{1,lp}(t,t)\right), \qquad (13)$$

where

$$\begin{aligned}\tilde{G}^<_{lp,1}(t,t) = \ & \int dt_1 g^<_{lp}(t,t_1)\tau_p^l \tilde{G}^A_{1,1}(t_1,t) + \\ & + \int dt_1 g^R_{lp}(t,t_1)\tau_p^l \tilde{G}^<_{1,1}(t_1,t), \\ \tilde{G}^<_{1,lp}(t,t) = \ & \int dt_1 \tilde{G}^<_{1,1}(t,t_1)\tau_p^l g^A_{lp}(t_1,t) + \\ & + \int dt_1 \tilde{G}^R_{1,1}(t,t_1)\tau_p^l g^<_{lp}(t_1,t).\end{aligned}$$

The parameter p corresponds to the density of states inside both reservoirs, $g_{\alpha p}(\omega)$ is the Green's function of reservoir α when it is disconnected from the chain, where α takes values l and r, $G^R_{n,m}(t, t1)$ are the exact are the retarded and advanced Green's functions of the chain, accounting for tunneling transitions into the reservoirs.

Crucially, the tunneling Hamiltonian (11) and the tunneling current (12) are expressed in terms of real electron operators, and they directly provide the actual electric current in the system. It should be noted that attempts to use effective Hamiltonians in terms of Majorana quasiparticle operators often lead, in our opinion, to questionable results, as handling Majorana operators requires great caution and precision. Due to the Clifford algebra commutation relations, there is no Wick's theorem directly applicable to Majorana operators, and pair correlators do not have the meaning of Green's functions, which form the basis of conventional diagrammatic techniques. In the calculations presented in this paper, we do not encounter any difficulties that we would have faced if we had worked with Majorana operators. For the problem of a finite Kitaev chain of arbitrary length, inserted between two leads and described by Hamiltonian (11), we have exactly computed the electronic current

(12). It is not surprising that some discrepancies may arise between our results and those of [26, 27, 25] and other authors, as the latter were obtained using a number of approximations in the Majorana operator representation.

In what follows, we assume, as usual, that due to the large number of particles and degrees of freedom in each reservoir, the particle distribution function does not significantly change throughout the experiment, and thus each reservoir remains practically in equilibrium. However, the system as a whole is not in equilibrium, although in this section, we consider it stationary, meaning the current does not change over time. Thus, Equation (13) can be rewritten using frequency-dependent Green's functions as follows:

$$I_l = -\sum_p \tau_p^l \int \frac{d\omega}{2\pi}\left(\tilde{G}^{<}_{lp,1}(\omega) - \tilde{G}^{<}_{1,lp}(\omega)\right). \quad (14)$$

where

$$\begin{aligned} \tilde{G}^{<}_{lp,1}(\omega) &= g^{<}_{lp}(\omega)\tau_p^l \tilde{G}^{A}_{1,1}(\omega) + g^{R}_{lp}(\omega)\tau_p^l \tilde{G}^{<}_{1,1}(\omega), \\ \tilde{G}^{<}_{1,lp}(t,t) &= \tilde{G}^{<}_{1,1}(\omega)\tau_p^l g^{A}_{lp}(\omega) + \tilde{G}^{R}_{1,1}(\omega)\tau_p^l g^{<}_{lp}(\omega). \end{aligned} \quad (15)$$

We can simplify this expression by introducing the irreducible part

$$\Sigma_\alpha^{R(A,<)}(\omega) = \sum_p \left(\tau_p^\alpha\right)^2 g_{\alpha p}^{R(A,<)}(\omega). \quad (16)$$

Then, we can use the identity

$$\Sigma_\alpha^{<}(\omega) = n_\alpha(\omega)\left(\Sigma_\alpha^{A}(\omega) - \Sigma_\alpha^{R}(\omega)\right),$$

where $n_\alpha(\omega)$ are the Fermi-Dirac distribution functions for the l and r reservoirs.

Thus, Equation (14) can be rewritten as:

$$\begin{aligned} \hat{I}_l = & -\int \frac{d\omega}{2\pi}\left(\Sigma_l^{A}(\omega) - \Sigma_l^{R}(\omega)\right)\times \\ & \times\left(n_l(\omega)\left(\tilde{\Gamma}^{A}_{1,1}(\omega) - \tilde{\Gamma}^{R}_{1,1}(\omega)\right) - \tilde{\Gamma}^{<}_{1,1}(\omega)\right). \end{aligned} \quad (17)$$

Here, the current I_l is determined by the upper-left element of the matrix $\hat{I}_l$ ($\hat{I}_l^{11}$).

An expression of this type in terms of nonequilibrium Green's functions was first derived in [24] and later applied in [28]. At first glance, this expression appears asymmetric with respect to the left and right contacts. However, in the stationary case, a properly calculated current (17) can always be rewritten in an explicitly symmetric form.

In our case, Equation (17) can be further simplified using the relations

$$\begin{aligned} \tilde{\Gamma}^{<}_{1,1}(\omega) &= \tilde{\Gamma}^{R}_{1,1}(\omega)\Sigma_l^{<}(\omega)\tilde{\Gamma}^{A}_{1,1}(\omega) + \\ &+ \tilde{\Gamma}^{R}_{1,N}(\omega)\Sigma_r^{<}(\omega)\tilde{\Gamma}^{A}_{N,1}(\omega), \end{aligned} \quad (18)$$

$$\begin{aligned} \tilde{\Gamma}^{\delta}_{1,1}(\omega) &= \tilde{\Gamma}^{R}_{1,1}(\omega)\Sigma_l^{\delta}(\omega)\tilde{\Gamma}^{A}_{1,1}(\omega) + \\ &+ \tilde{\Gamma}^{R}_{1,N}(\omega)\Sigma_r^{\delta}(\omega)\tilde{\Gamma}^{A}_{N,1}(\omega). \end{aligned} \quad (19)$$

where

$$\tilde{\Gamma}^{\delta}_{1,1}(\omega) = \tilde{\Gamma}^{A}_{1,1}(\omega) - \tilde{\Gamma}^{R}_{1,1}(\omega),$$

$$\Sigma_\alpha^{\delta}(\omega) = \Sigma_\alpha^{A}(\omega) - \Sigma_\alpha^{R}(\omega).$$

Using the wide-band approximation for the reservoirs, we assume that for the considered values of ω, the condition $\Sigma^A_{l(r)}(\omega) \approx i\gamma_{l(r)}$, $\Sigma^R_{l(r)}(\omega) \approx -i\gamma_{l(r)}$, holds, where $\gamma_{l(r)} = \pi\nu^{l(r)}(\tau_p^{l(r)})^2$ and $\nu^{l(r)}$ are the densities of states in the rese rvoirs $l(r)$.

Direct substitution gives:

$$\hat{I}_l = 4\gamma_l\gamma_r \int \frac{d\omega}{2\pi}\tilde{\Gamma}^{R}_{1,N}(\omega)\tilde{\Gamma}^{A}_{N,1}(\omega)\left(n_l(\omega) - n_r(\omega)\right). \quad (20)$$

A formula of this type was derived in [24]. It should be noted that the obtained equation for the current through the system is symmetric with respect to its two edges. Naturally, this implies that, in the stationary case, the current flowing into the system equals the current flowing out of it. The conservation of total current cannot be violated in any system and does not require additional conditions, such as equal tunneling rates or symmetrically applied voltages at different edges. Thus, the appearance of asymmetric expressions for stationary tunneling current, as obtained in some works on Kitaev chain-type systems (e.g., [29]), signals the need to verify the applied approximations. This statement remains valid even for interacting systems, but deriving an explicitly symmetric expression in such cases is more challenging. Examples of such calculations for systems with electron-phonon interactions can be found, for example, in [30, 31]. We emphasize that Equation (20) is exact and explicitly symmetric for the left and right contacts.

Since we aim to study the low-energy bound state corresponding to the "Majorana mode", we consider the case where the applied voltage is smaller than the superconducting gap. In this case, we exclude the influence of quasiparticle states from the continuous

spectrum. To express $\tilde{\Gamma}^R_{1,N}(\omega)$ through the Green's functions of the isolated chain $\Gamma^R_{n,m}(\omega)$, we use Dyson's equation:

$$\begin{aligned}\tilde{\Gamma}^R_{n,m}(\omega) = \ & \Gamma^R_{n,m}(\omega) + \Gamma^R_{n,1}(\omega)\Sigma^R_l(\omega)\tilde{\Gamma}^R_{1,m}(\omega) \\ & +\Gamma^R_{n,N}(\omega)\Sigma^R_r(\omega)\tilde{\Gamma}^R_{N,m}(\omega).\end{aligned} \tag{21}$$

Simple algebraic transformations yield:

$$\begin{aligned}\tilde{\Gamma}^R_{1,N}(\omega) = & \left[\hat{\mathrm{I}} + \gamma_l\gamma_r\left(\hat{\mathrm{I}} + i\gamma_l\Gamma^R_{1,1}(\omega)\right)\Gamma^R_{1,N}(\omega)\times\right. \\ & \left.\times\left(\hat{\mathrm{I}} + i\gamma_r\Gamma^R_{N,N}(\omega)\right)\Gamma^R_{N,1}(\omega)\right]^{-1}\times \\ & \times\left(\hat{\mathrm{I}} + i\gamma_l\Gamma^R_{1,1}(\omega)\right)\Gamma^R_{1,N}(\omega)\left(\hat{\mathrm{I}} + i\gamma_r\Gamma^R_{N,N}(\omega)\right).\end{aligned}$$

where $\hat{\mathrm{I}}$ is the identity matrix. The explicit form of the Green's functions $\Gamma^R_{n,m}(\omega)$ for $|\omega| \ll |\Delta|, t$ is derived in the Appendix. A simpler form can be obtained for $\Delta^2/(t\gamma) \gg 1$. Retaining the leading terms in (55) for this parameter, we get:

$$\tilde{\Gamma}^R_{1,N}(\omega) = \frac{C\omega_0}{\omega^2 - \omega_0^2 + 2i(\gamma_l + \gamma_r)C\omega - 4\gamma_l\gamma_r C^2}\begin{pmatrix} 1 & -\dfrac{\Delta}{|\Delta|} \\ \dfrac{\Delta^*}{|\Delta|} & -1 \end{pmatrix}, \tag{22}$$

$$C \equiv -\frac{|\Delta|(4t^2 - \mu^2)}{2t(4(t^2 - |\Delta|^2) - \mu^2)}(\chi_+ - \chi_-)^2 = \frac{|\Delta|(4t^2 - \mu^2)}{2t(t + |\Delta|)^2}. \tag{23}$$

Substituting this result into (20), we obtain:

$$I_l = \int\frac{d\omega}{2\pi}\frac{8\gamma_l\gamma_r C^2\omega_0^2}{\left|\omega^2 - \omega_0^2 + 2i(\gamma_l + \gamma_r)C\omega - 4\gamma_l\gamma_r C^2\right|^2}(n_l(\omega) - n_r(\omega)). \tag{24}$$

These and further calculations are performed for the following parameter hierarchy: $t > \Delta > \gamma_{l,r}$. For the case $\gamma_{l,r} > \Delta$, we cannot exclude the influence of the continuous part of the spectrum on conductivity, and information about low-energy resonances is lost, so this case is not considered here.

We see that the magnitude of the current (25) is directly proportional to ω_0^2, meaning that the current decreases exponentially with increasing chain length. Moreover, if $\omega_0 = 0$, which is typically associated with Majorana particles, then no current flows through the system at all. Note that Equation (25) is symmetric with respect to the contact parameters *l* and *r*, as expected. A similar expression for the normal component of the current was obtained in the quasiclassical approach in [18], where it was also noted that the zero-bias peak in tunneling conductance is unlikely to be observed for a realistic ratio between ω_0 and $\gamma_{l,r}$.

The tunneling conductance peak associated with Majorana states was also studied in [32]. That study considered a single NS contact, where it was assumed that the chemical potential of the superconductor was somehow fixed. The problem was solved using the effective transmission coefficient method for quasiparticles, which, in the presence of a bound state, always leads to formulas of type (25). However, the peak amplitude for the two different systems – a single NS contact and a superconductor between two normal contacts – cannot be directly compared due to the problem of fixing the superconducting chemical potential. It is worth noting that results similar to those in [32] for the current in an NS contact, considering Majorana states, can also be obtained using the methods from [33].

If in Equation (25) the applied voltage is greater than the width of localized states, but less than the superconducting gap, meaning $n_l(\omega) - n_r(\omega) = 1$ for $|\omega| \lesssim \gamma_l, \gamma_r$ then we obtain a simple final expression for the tunneling current associated with Majorana modes:

$$I_l = \frac{2\gamma_l\gamma_r C\omega_0^2}{\gamma_l + \gamma_r}\frac{1}{4\gamma_l\gamma_r C^2 + \omega_0^2}. \tag{25}$$

Thus, the magnitude of the current is always determined by the smallest transfer rate present in the system (the weakest link); in our case, these rates are defined by the parameters $\omega_0^2/(\gamma_l+\gamma_r), \gamma_l, \gamma_r$. If $\omega_0^2 \geq C^2\gamma_l\gamma_r$, then the general equation (25) leads to a current proportional to $\gamma_l\gamma_r/(\gamma_l+\gamma_r)$, which is the usual expression for tunneling through an intermediate state. For the considered system, the physically reasonable relation is $\omega_0 \ll \gamma_l, \gamma_r$. Using this, we obtain:

$$I_l = \frac{\omega_0^2}{2C(\gamma_l+\gamma_r)}. \tag{26}$$

If we use Equations (8) and (23), this formula gives:

$$I = \begin{cases} \dfrac{2\Delta t}{(\gamma_l+\gamma_r)} e^{-2N(\Delta/t)}, \Delta \ll t, \\ \dfrac{2t^2}{(\gamma_l+\gamma_r)} e^{-N\ln(2t/(t-\Delta))}, (t-\Delta) \ll t. \end{cases} \tag{27}$$

For arbitrary parameters $\mu < \Delta < t$, the current is always small in long chains. In the case of $\omega_0 = 0$, which is considered the most favorable scenario for observing unusual topological properties, we will not be able to observe a zero-bias peak in the tunneling conductance at all. This observation holds true for the model considered in this paper, where the chain has two contacts at its edges. In any experiment measuring stationary current, at least two external leads are required, connected to the "left" and "right" edges of the system. Of course, there are more complex multi-contact configurations, but their analysis is beyond the scope of this paper. Real hybrid semiconductor-superconductor structures, which simulate the Kitaev chain, require the consideration of a model Hamiltonian that describes a semiconductor nanowire with strong spin-orbit interaction, which is coupled due to the proximity effect to an underlying superconducting layer. In this case, the superconductor can be considered as a reservoir with a fixed chemical potential, and the "second contact" as the interface between the semiconductor and the superconductor. Alternatively, instead of edge connections, we could also consider a Kitaev chain lying on a substrate, where all chain atoms are weakly coupled to corresponding substrate atoms. In this case, the "second contact" with the reservoir becomes spatially distributed. This problem can be solved, but it is different from the one considered in this paper. Nevertheless, if the overlap of the localized state with the reservoir states is small, then the zero-bias current peak should also be small. Its magnitude in the case of a spatially distributed "second contact" will not decay exponentially with chain length, but will still be much smaller than what would be expected from naive formulas. This may be a possible reason why the zero-bias peak is often poorly observed in conventional tunneling experiments [14].

We want to emphasize that naively applied general formulas for the tunneling current between two contacts often lead to misleading results when used for low-dimensional systems, such as the Kitaev chain [21], due to the possible appearance of localized states in the contact region.

The lowest-order response (second order in the tunneling coupling) of quantum mechanical perturbation theory describes the current only at the initial moment after the tunneling connection is "switched on". However, the stationary tunneling current can only be calculated using the full system of kinetic equations, or equivalently, the full system of equations for the nonstationary Keldysh-Green's functions. Only in simple systems with a continuous spectrum, where rapid electron relaxation to equilibrium is implicitly assumed, is the formula based on the equilibrium local density of states of the leads guaranteed to be valid.

To clarify this idea, let us consider a tunneling contact with a localized state at the edge of one of the leads. This localized state creates a sharp peak in the local density of states and contributes to the simplest formula for tunneling current. Suppose this state is empty at the initial moment (i.e., lies above the Fermi level). Then, immediately after applying a positive bias voltage to the other lead, the current begins to flow into this empty localized state. However, after some relaxation time, determined by the tunneling rate, this state becomes occupied, and from that point onward, no more electrons can tunnel into it. The stationary tunneling current then vanishes, even though the simplest formula still predicts a "zero-bias peak" in the tunneling conductance. For this localized state to contribute to the stationary current, some inelastic processes must be included, which are responsible for removing (or adding) electrons from this localized state. For a finite system, it is also possible that this localized state at one edge has some overlap with the second contact. (This corresponds to our case and the case of a distributed "grounded contact" in a real system.)

In the usual formula for tunneling current, which relies on the local density of states of the contacts, it is

implicitly assumed that at any moment, the chemical potentials of all contact states are fixed. To maintain a constant chemical potential, the system must be connected to some reservoir via a contact that allows for particle exchange. Thus, when we say that we fix the chemical potential of localized states, we are implicitly including some inelastic relaxation processes or a direct connection to a reservoir for these states.

4. NONSTATIONARY CURRENT

Now, let us attempt to answer the question of what the typical time scales are for current or charge transfer from one edge of the chain to the other. We will pose the problem differently than in [18], where the effect of periodic modulation of tunnel barrier transparency on zero-bias tunneling conductance was studied. An interesting result in that study was the discovery and analysis of resonance between the external driving frequency and the splitting of Majorana states ω_0. In our case, we are interested in the characteristic speeds of transient processes. To do this, let us assume that the system is initially in equilibrium at $t < 0$, and then at $t = 0$, a voltage is applied to one of the leads. This additional voltage induces a nonstationary current, which at t→∞ reaches the stationary value (25).

The applied voltage shifts the energy levels in the reservoirs by V_α, where the index α denotes the reservoir. Thus, the reservoir Hamiltonian can now be written as

$$\begin{aligned}\widehat{H}_\alpha(t) = &\sum_p \tau_p^\alpha \left(h_p^{\alpha+}\psi_1 + \psi_1^+ h_p^\alpha\right) + \\ &+\sum_p \left(E_p^\alpha + V_\alpha\theta(t)\right)h_p^{\alpha+}h_p^\alpha.\end{aligned} \tag{28}$$

The current flowing from the left reservoir into the system is given by (for the "right" contact r, all formulas can be written similarly):

$$\begin{aligned}I_l(t) = -\int dt' \quad &\left(\Sigma_l^<(t,t')\widetilde{G}^A_{1,1}(t',t) + \right. \\ &+\Sigma_l^R(t,t_1)\widetilde{G}^<_{1,1}(t',t) - \\ &-\widetilde{G}^<_{1,1}(t,t')\Sigma_l^A(t',t) - \\ &\left.-\widetilde{G}^R_{1,1}(t,t')\Sigma_l^<(t',t)\right).\end{aligned} \tag{29}$$

Here, the irreducible part takes the form:

$$\begin{aligned}\Sigma_\alpha^R(t,t') = &-i\sum_p(\tau_p^\alpha)^2\theta(t-t')\times \\ &\times\exp\left(-iE_p^\alpha(t-t') - iV_\alpha\int_{t'}^t dt_1\theta(t_1)\right),\end{aligned} \tag{30}$$

$$\begin{aligned}\Sigma_\alpha^<(t,t') = &i\sum_p(\tau_p^\alpha)^2 n_p^\alpha \times \\ &\times\exp\left(-iE_p^\alpha(t-t') - iV_\alpha\int_{t'}^t dt_1\theta(t_1)\right).\end{aligned} \tag{31}$$

In the frequency representation, these expressions correspond to the following formulas:

$$\begin{aligned}\Sigma_\alpha^R(\omega,\omega') = -i(\tau^\alpha)^2\int d\varepsilon\nu^\alpha(\varepsilon) \quad &\left[-\frac{1}{\omega-\varepsilon-V_\alpha+2i\delta}\left(-\frac{1}{\omega'-\omega-2i\delta}+\frac{1}{\omega'-\omega+V_\alpha}\right)+\right. \\ &\left.+\frac{1}{\omega'-\varepsilon+2i\delta}\left(-\frac{1}{\omega-\omega'-V_\alpha}+\frac{1}{\omega-\omega'-2i\delta}\right)\right],\end{aligned} \tag{32}$$

$$\Sigma_\alpha^<(\omega,\omega') = i(\tau^\alpha)^2\int \quad d\varepsilon\nu^\alpha(\varepsilon)n^\alpha(\varepsilon)\left(\frac{1}{\omega-\varepsilon-V_\alpha+i\delta}-\frac{1}{\omega-\varepsilon-i\delta}\right)\left(\frac{1}{\omega'-\varepsilon-V_\alpha-i\delta}-\frac{1}{\omega'-\varepsilon+i\delta}\right). \tag{33}$$

where $\nu^\alpha(\varepsilon)$ is the density of states in the reservoir α, $\delta \to +0$. For simplicity, we assume that τ^α does not depend on p. In the wide-band approximation, where we assume that $\nu(\varepsilon)$ remains constant for $\varepsilon \sim \omega,\omega',V_{l,r}$, these expressions simplify to:

$$\begin{aligned}\Sigma_l^R(\omega,\omega') = &-i\gamma_l 2\pi\delta(\omega'-\omega), \\ \Sigma_l^<(\omega,\omega') = &\frac{i\gamma_l}{\pi}\int d\varepsilon n^l(\varepsilon)\times \\ &\times\left(\frac{1}{\omega-\varepsilon-V_l+i\delta}-\frac{1}{\omega-\varepsilon-i\delta}\right)\times \\ &\times\left(\frac{1}{\omega'-\varepsilon-V_l-i\delta}-\frac{1}{\omega'-\varepsilon+i\delta}\right).\end{aligned}$$

As a result, in the frequency representation, Equation (29) simplifies to:

$$\begin{aligned}\hat{I}_l(\omega) = -\int\frac{d\Omega}{2\pi} \quad &\left(\Sigma_l^<(\Omega,\Omega-\omega)\widetilde{\Gamma}^\delta_{1,1}(\Omega-\omega) - \right. \\ &\left.-2i\gamma_l\widetilde{\Gamma}^<_{1,1}(\Omega,\Omega-\omega)\right).\end{aligned} \tag{34}$$

where

$$\begin{aligned}\tilde{\Gamma}^{<}_{1,1}(\Omega,\Omega-\omega) = &\quad \tilde{\Gamma}^{R}_{1,1}(\Omega)\Sigma^{<}_{l}(\Omega,\Omega-\omega)\tilde{\Gamma}^{A}_{1,1}(\Omega-\omega) + \\ + &\quad \tilde{\Gamma}^{R}_{1,N}(\Omega)\Sigma^{<}_{r}(\Omega,\Omega-\omega)\tilde{\Gamma}^{A}_{N,1}(\Omega-\omega), \\ \tilde{\Gamma}^{\delta}_{1,1}(\Omega-\omega) = &\quad \tilde{\Gamma}^{A}_{1,1}(\Omega-\omega) - \tilde{\Gamma}^{R}_{1,1}(\Omega).\end{aligned}$$

Using Dyson's equations for retarded and advanced Green's functions, we can show that:

$$\begin{aligned}\tilde{\Gamma}^{\delta}_{1,1}(\Omega-\omega) = &\quad \omega\sum_{n=1}^{N}\Gamma^{R}_{1,n}(\Omega)\Gamma^{A}_{n,1}(\Omega-\omega) + \\ &+\tilde{\Gamma}^{R}_{1,1}(\Omega)2i\gamma_{l}\tilde{\Gamma}^{A}_{1,1}(\Omega-\omega) + \\ &+\tilde{\Gamma}^{R}_{1,N}(\Omega)2i\gamma_{r}\tilde{\Gamma}^{A}_{N,1}(\Omega-\omega).\end{aligned}$$

Substituting these last expressions into (34), we obtain:

$$\begin{aligned}\hat{I}_{l}(\omega) = -\int\frac{d\Omega}{2\pi}\Bigg(&\omega\sum_{n=1}^{N}\tilde{\Gamma}^{R}_{1,n}(\Omega)\tilde{\Gamma}^{A}_{n,1}(\Omega-\omega) \\ &\times\Sigma^{<}_{l}(\Omega,\Omega-\omega) \\ +2i\Big(&\gamma_{r}\Sigma^{<}_{l}(\Omega,\Omega-\omega) - \gamma_{l}\Sigma^{<}_{r}(\Omega,\Omega-\omega)\times\Big) \\ &\times\tilde{\Gamma}^{R}_{1,N}(\Omega)\tilde{\Gamma}^{A}_{N,1}(\Omega-\omega)\Bigg).\end{aligned} \tag{35}$$

Here

$$\begin{aligned}\Sigma^{<}_{r}(\omega,\omega') &= \frac{i\gamma_{r}}{\pi}\int d\varepsilon n^{l}(\varepsilon)\times \\ &\times\left(\frac{1}{\omega-\varepsilon-V_{r}+i\delta}-\frac{1}{\omega-\varepsilon-i\delta}\right)\times \\ &\times\left(\frac{1}{\omega'-\varepsilon-V_{r}-i\delta}-\frac{1}{\omega'-\varepsilon+i\delta}\right).\end{aligned} \tag{36}$$

We see that the first term in (35) exists only if $V_l \neq 0$ and does not directly depend on the properties of the right reservoir *r*.

This means that this term corresponds to the filling of states at the left edge of the chain due to a change in its chemical potential.

Consequently, the second term represents the current that flows from one reservoir to another through the entire chain.

If we consider only the second term, we obtain:

$$\begin{aligned}\hat{I}_{l}(t) = &\quad \frac{2\gamma_{l}\gamma_{r}}{\pi}\int d\varepsilon dV\,\widehat{M}_{1,N}(t,\varepsilon,V)\left(\widehat{M}_{1,N}(t,\varepsilon,V)\right)^{*}\times \\ &\times\left[n^{l}(\varepsilon)\delta(V-V_{l}) - n^{r}(\varepsilon)\delta(V-V_{r})\right],\end{aligned}$$

$$\begin{aligned}\widehat{M}_{1,N}(t,\varepsilon,V) = &\quad \int\frac{d\Omega}{2\pi}e^{-i\Omega t}\tilde{\Gamma}^{R}_{1,N}(\Omega)\times \\ &\times\left(\frac{1}{\Omega-\varepsilon-V+i\delta}-\frac{1}{\Omega-\varepsilon-i\delta}\right).\end{aligned}$$

Since our goal is to study the propagation of perturbations through the chain, we assume that at time $t = 0$, the voltage changes only at the right contact, and we observe the time-dependent current at the left contact under the condition $V_l = 0$. Then, by direct calculations, we obtain that:

$$\begin{aligned}\widehat{M}_{1,N}(t,\varepsilon,V) = - &\quad i\theta(-t)\frac{e^{-i\varepsilon t}C\,\omega_{0}}{(\varepsilon+iC(\gamma_{l}+\gamma_{r}))^{2}-\bar{\omega}^{2}}\begin{pmatrix}1 & -\dfrac{\Delta}{|\Delta|}\\ \dfrac{\Delta^{*}}{|\Delta|} & -1\end{pmatrix} - i\theta(t)\frac{e^{-i(\varepsilon+V_{r})t}C\,\omega_{0}}{(\varepsilon+V+iC(\gamma_{l}+\gamma_{r}))^{2}-\bar{\omega}^{2}}\begin{pmatrix}1 & -\dfrac{\Delta}{|\Delta|}\\ \dfrac{\Delta^{*}}{|\Delta|} & -1\end{pmatrix} - \\ &-\frac{iC\,\omega_{0}\theta(t)}{2\bar{\omega}}e^{-C(\gamma_{l}+\gamma_{r})t-i\bar{\omega}t}\begin{pmatrix}1 & -\dfrac{\Delta}{|\Delta|}\\ \dfrac{\Delta^{*}}{|\Delta|} & -1\end{pmatrix}\left(-\frac{1}{\varepsilon+V+iC(\gamma_{l}+\gamma_{r})-\bar{\omega}}+\frac{1}{\varepsilon+iC(\gamma_{l}+\gamma_{r})-\bar{\omega}}\right) - \\ &-\frac{iC\,\omega_{0}\theta(t)}{-2\bar{\omega}}e^{-C(\gamma_{l}+\gamma_{r})t+i\bar{\omega}t}\begin{pmatrix}1 & -\dfrac{\Delta}{|\Delta|}\\ \dfrac{\Delta^{*}}{|\Delta|} & -1\end{pmatrix}\left(-\frac{1}{\varepsilon+V+iC(\gamma_{l}+\gamma_{r})+\bar{\omega}}+\frac{1}{\varepsilon+iC(\gamma_{l}+\gamma_{r})+\bar{\omega}}\right).\end{aligned}$$

$$I_l(t) = \frac{4\gamma_l\gamma_r C^2\omega_0^2}{\pi}\theta(t)\int d\varepsilon n^l(\varepsilon)\left|\frac{1}{(\varepsilon + iC(\gamma_l+\gamma_r))^2 - \bar{\omega}^2}\right|^2 -$$
$$-\frac{4\gamma_l\gamma_r C^2\omega_0^2}{\pi}\theta(t)\int d\varepsilon n^r(\varepsilon - V_r)\times$$
$$\times\left|-\frac{e^{-i\varepsilon t}}{(\varepsilon + iC(\gamma_l+\gamma_r))^2 - \bar{\omega}^2} - \frac{1}{2\bar{\omega}}e^{-C(\gamma_l+\gamma_r)t - i\bar{\omega}t}\left(-\frac{1}{\varepsilon + iC(\gamma_l+\gamma_r) - \bar{\omega}} + \frac{1}{\varepsilon - V_r + iC(\gamma_l+\gamma_r) - \bar{\omega}}\right) + \right.$$
$$\left. + \frac{1}{2\bar{\omega}}e^{-C(\gamma_l+\gamma_r)t + i\bar{\omega}t}\left(-\frac{1}{\varepsilon + iC(\gamma_l+\gamma_r) + \bar{\omega}} + \frac{1}{\varepsilon - V_r + iC(\gamma_l+\gamma_r) + \bar{\omega}}\right)\right|^2.$$

Here

$$\bar{\omega} = \sqrt{\omega_0^2 - C^2(\gamma_l - \gamma_r)^2}. \tag{37}$$

As expected, if $t\to\infty$, the current approaches its stationary value (25):

$$I_l(t\to\infty) = \frac{4\gamma_l\gamma_r C^2\omega_0^2}{\pi}\theta(t)\times$$
$$\times\int d\varepsilon(n^l(\varepsilon) - n^r(\varepsilon - V_r))\times$$
$$\times\left|\frac{1}{(\varepsilon + iC(\gamma_l+\gamma_r))^2 - \bar{\omega}^2}\right|^2.$$

If $t \to +0$, the current at the opposite edge of the chain is not observed, illustrating the continuity of the current change when passing through $t = 0$:

$$I_l(t\to+0) = \frac{4\gamma_l\gamma_r C^2\omega_0^2}{\pi}\theta(t)\times$$
$$\times\int d\varepsilon(n^l(\varepsilon) - n^r(\varepsilon))\times$$
$$\times\left|\frac{1}{(\varepsilon + iC(\gamma_l+\gamma_r))^2 - \bar{\omega}^2}\right|^2 = 0.$$

If now, as in the previous section, we are interested in the role of "Majorana states," we apply an additional voltage to the right contact, which is greater than the width of the localized states but less than the value of the superconducting gap. This means that the conditions

$$n^l(\varepsilon) = n^r(\varepsilon) = 0, \quad n^r(\varepsilon - V_r) = 1.$$

are satisfied for $\varepsilon \lesssim \gamma_l, \gamma_r$

The current is defined as

$$I_l(t) = -\frac{2\gamma_l\gamma_r C\omega_0^2}{\gamma_l+\gamma_r}\frac{1}{\omega_0^2 + 4C^2\gamma_l\gamma_r}\theta(t) +$$
$$+\frac{2\gamma_l\gamma_r C\omega_0^2}{(\gamma_l+\gamma_r)\bar{\omega}^2}e^{-2C(\gamma_l+\gamma_r)t}\theta(t) -$$
$$-i\gamma_l\gamma_r\frac{C^2\omega_0^2}{\bar{\omega}^2}\frac{e^{-2C(\gamma_l+\gamma_r)t + 2i\bar{\omega}t}}{\bar{\omega} + iC(\gamma_l+\gamma_r)}\theta(t) + \tag{38}$$
$$+i\gamma_l\gamma_r\frac{C^2\omega_0^2}{\bar{\omega}^2}\frac{e^{-2C(\gamma_l+\gamma_r)t - 2i\bar{\omega}t}}{\bar{\omega} - iC(\gamma_l+\gamma_r)}\theta(t).$$

We consider the case $\gamma_r, \gamma_l \gg \omega_0$ under the assumption that ω_0 is always small. However, for very symmetric tunneling coupling with the leads, we could have $\omega_0^2 \gg (\gamma_r - \gamma_l)^2$. This case appears unrealistic, but it demonstrates an oscillating current signal at the left edge:

$$I_l(t) = -\frac{\omega_0^2}{2C(\gamma_l+\gamma_r)}\times$$
$$\times\left[1 - e^{-2C(\gamma_l+\gamma_r)t}\right] -$$
$$-\left[\frac{\omega_0}{2}\sin(2\omega_0 t) - \frac{C(\gamma_l+\gamma_r)}{2}(1 - \cos(2\omega_0 t))\right]$$
$$\times e^{-2C(\gamma_l+\gamma_r)t}. \tag{39}$$

If $\omega_0 \ll |\gamma_r - \gamma_l|$ and $t > 0$, Equation (38) simplifies to:

$$I_l(t) = -\frac{\omega_0^2}{2C(\gamma_l+\gamma_r)}\times$$
$$\times\left[1 + \frac{4\gamma_l\gamma_r}{(\gamma_l-\gamma_r)^2}e^{-2C(\gamma_l+\gamma_r)t}\right. \tag{40}$$
$$\left. -\frac{(\gamma_l+\gamma_r)}{(\gamma_l-\gamma_r)^2}\left(\gamma_l e^{-4C\gamma_r t} + \gamma_r e^{-4C\gamma_l t}\right)\right].$$

Note that the negative sign indicates that the current flows from *r* to *l*. For significantly different tunneling rates, for example, $\gamma_r \gg \gamma_l$, the time evolution of the leading contribution to the current is determined by the slowest rate:

$$I_l(t) = -\frac{\omega_0^2}{2C\gamma_r}\left[1 - e^{-4C\gamma_l t}\right]. \quad (41)$$

The final formula shows that if $\gamma_l \to 0$, the current signal at the other end of the chain increases very slowly.

5. CONCLUSION

This paper demonstrates that the transport properties of a finite-length Kitaev chain can be fully investigated using the conventional Green's function technique. For any nonstationary problem, this formalism appears much more convenient than the language of Majorana fermions or other methods, allowing for the exact analytical results. Our calculations bridge the gap between phenomenological parameters for quasiparticles in quasiclassical calculations and the microscopic description of quasi-one-dimensional superconductors.

It has been shown that the stationary tunneling current through a finite chain is always determined by the lowest transfer rate among the parameters $\omega_0^2/(\gamma_l+\gamma_r), \gamma_l, \gamma_r$, provided the applied voltage is less than the superconducting gap. For arbitrary $\mu < |\Delta| < t$, the stationary current is always exponentially small for long chains. It should be noted that for a finite Kitaev chain placed between two external thermostat contacts, no significant peak can be observed at ω_0 in the tunneling conductance. Furthermore, in the case of $\omega_0 = 0$, the stationary current completely vanishes.

We have also obtained the time-dependent behavior of the tunneling current following a sudden change in the bias voltage at one of the leads. It was shown that the typical timescales of tunneling current evolution are primarily determined by the tunneling rates γ_l, γ_r from the left and right edge sites of the chain to the corresponding leads. Although the results presented here are for an ideal system, we can be confident – based on the conclusions of [34, 35]— that weak disorder does not significantly affect the properties of the ideal Kitaev chain. Therefore, only strong disorder can completely alter our results.

In conclusion, it is worth noting that when considering systems of multiple Kitaev chains, an effective description based on Coulomb blockade effects is often constructed. However, such an effective description is sensitive to charge transfer rates, which may be important for modern proposals related to signal transmission, quantum information exchange, and storage using Kitaev chains.

FUNDING

This work was supported by the Russian Science Foundation, Grant No. 23-22-00289.

APPENDIX ANALYTICAL DESCRIPTION OF THE ISOLATED KITAEV CHAIN

In this section, we present the formulas for the Green's functions of the isolated Kitaev chain.

As shown in [20], the exact solution for the Green's functions of the infinite chain can be written as:

$$\Gamma_{nm}^{0R}(\omega) = -\frac{1}{4(|\Delta|^2 - t^2)(A_+ - A_-)} \times$$

$$\times\left[\chi_+^{|n-m|}\widehat{M}_1 - \chi_-^{|n-m|}\widehat{M}_2\right]. \quad (42)$$

Here

$$\widehat{M}_1 = \begin{pmatrix} \dfrac{\omega - \mu - 2tA_+}{\sqrt{A_+^2 - 1}} & 2\Delta\,\mathrm{sign}(n-m) \\ -2\Delta^*\,\mathrm{sign}(n-m) & \dfrac{\omega + \mu + 2tA_+}{\sqrt{A_+^2 - 1}} \end{pmatrix},$$

$$\widehat{M}_2 = \begin{pmatrix} \dfrac{\omega - \mu - 2tA_-}{\sqrt{A_-^2 - 1}} & 2\Delta\,\mathrm{sign}(n-m) \\ -2\Delta^*\,\mathrm{sign}(n-m) & \dfrac{\omega + \mu + 2tA_-}{\sqrt{A_-^2 - 1}} \end{pmatrix}.$$

The complex value of the square root $\sqrt{A_\pm^2 - 1}$ is defined such that it has a branch cut along the interval $\sqrt{A_\pm^2 - 1}$ and takes positive values when $A_\pm > 1$.

$$A_\pm = \frac{t\mu \pm |\Delta|\sqrt{\mu^2 + 4(|\Delta|^2 - t^2)\left(1 - \dfrac{(\omega + i\delta)^2}{4|\Delta|^2}\right)}}{2(|\Delta|^2 - t^2)}, \quad (43)$$

$$\chi_{\pm} = A_{\pm} - \sqrt{A_{\pm}^2 - 1}. \quad (44)$$

We assume $\delta \to +0$. The Green's function for the Hamiltonian (2) can be written in terms of the Green's function of the infinite chain, using Dyson's equation with the perturbation $\hat{V}$,

$$\Gamma^R_{nm}(\omega) = \Gamma^{0R}_{nm}(\omega) + \Gamma^{0R}_{n0}(\omega) U \sigma_z \Gamma^R_{0m}(\omega) + \\ + \Gamma^{0R}_{n,N+1}(\omega) U \sigma_z \Gamma^R_{N+1,m}(\omega). \quad (45)$$

If we solve Equation (45) for $\Gamma^R_{nm}(\omega)$ and take the limit $U \to \infty$, we can find the exact solution for the Green's functions $\Gamma^R_{nm}(\omega)$:

$$\begin{aligned}\Gamma^R_{nm}(\omega) = \quad & \Gamma^{0R}_{nm}(\omega) - \Gamma^{0R}_{n0} \times \\ & \times(\omega)\left(\Gamma^{0R}_{0,0}(\omega) - \Gamma^{0R}_{0,N+1}(\omega)(\Gamma^{0R}_{N+1,N+1}(\omega))^{-1}\Gamma^{0R}_{N+1,0}(\omega)\right)^{-1} \times \\ & \times\left(\Gamma^{0R}_{0,m}(\omega) - \Gamma^{0R}_{0,N+1}(\omega)(\Gamma^{0R}_{N+1,N+1}(\omega))^{-1}\Gamma^{0R}_{N+1,m}(\omega)\right) - \\ & -\Gamma^{0R}_{n,N+1}(\omega)\left(\Gamma^{0R}_{N+1,N+1}(\omega) - \Gamma^{0R}_{N+1,0}(\omega)(\Gamma^{0R}_{00}(\omega))^{-1}\Gamma^{0R}_{0,N+1}(\omega)\right)^{-1} \times \\ & \times\left(\Gamma^{0R}_{N+1,m}(\omega) - \Gamma^{0R}_{N+1,0}(\omega)(\Gamma^{0R}_{00}(\omega))^{-1}\Gamma^{0R}_{0m}(\omega)\right).\end{aligned} \quad (46)$$

The matrix elements of $\Gamma^R_{nm}(\omega)$ describe the Green's functions of the finite chain, provided the indices satisfy the condition $0 < n,m < N+1$. It can be directly verified that $\Gamma^R_{nm}(\omega) = 0$ if one of the arguments n or m is positive, while the other is negative, giving us direct proof that our procedure effectively removes the site $n = 0$ from the system. The same is true for the site $n = N+1$.

We can see that the function $\Gamma^R_{nm}(\omega)$ may have a set of poles at values ω determined by the equation:

$$\det\left(\Gamma^{0R}_{00}(\omega) - \Gamma^{0R}_{N+1,0}(\omega)(\Gamma^{0R}_{00}(\omega))^{-1}\Gamma^{0R}_{0,N+1}(\omega)\right) = 0. \quad (47)$$

Since $\Gamma^{0R}_{nm}(\omega)$ has no poles inside the superconducting gap, it can be assumed that the solutions of this equation correspond to the energies of states localized at the chain edges. Direct substitution of the Green's functions (42) allows us to find the solution for ω at arbitrary parameter values.

For the semi-infinite chain, if $N \to \infty$, the situation simplifies significantly. Equation (47) simplifies to:

$$\det\left(\Gamma^{0R}_{0,0}(\omega)\right) = 0, \quad (48)$$

and it has only one solution in the gap $\omega = 0$. This solution does not arise if $|\mu| > 2t$. This pole at $\omega = 0$ exists in the Green's function $\Gamma^R_{nm}(\omega)$ only if both n and m are positive or both are negative, for any set of parameters t, μ, Δ satisfying the condition

$$t^2 > (\mu / 2)^2 + \Delta^2,$$

the condition that separates the topologically nontrivial and trivial phases. This means that the system described by the Hamiltonian (2) has two states with energy $\omega = 0$: one to the left and one to the right of the defect, which cuts the chain into two subsystems.

If we now consider a long finite chain of length T, we can write the equation for localized states as

$$\det\left[\Gamma^{(X)R}_{N+1,N+1}(\omega)\right] = 0, \quad (49)$$

where $\Gamma^{(X)R}_{N+1,N+1}(\omega)$ is the Green's function for the semi-infinite chain:

$$\Gamma^{(X)R}_{nm}(\omega) \equiv \\ \equiv \Gamma^{0R}_{nm}(\omega) - \Gamma^{0R}_{n0}(\omega)\left(\Gamma^{0R}_{0,0}(\omega)\right)^{-1}\Gamma^{0R}_{0,m}(\omega). \quad (50)$$

Since we are interested in bound states within the gap with energies close to zero, the calculations can be simplified using the following fact. For $\omega \to 0$, the values of χ satisfy the condition $|\chi_{\pm}| < 1$. Indeed, for $\omega = 0$, Equation (44) gives

$$\chi_{\pm} = \frac{-\mu \pm i\sqrt{4t^2 - \left(\mu^2 + 4\,|\,\Delta\,|^2\right)}}{2(t + |\Delta|)}. \quad (51)$$

As a result,

$$|\chi|^2 = \left(\frac{t - |\Delta|}{t + |\Delta|}\right). \quad (52)$$

This means that $|\chi| < 1$ for $t^2 > ((\mu/2)^2 + \Delta^2)$ and $\omega \ll |\Delta|$. Thus, quantities like $|\chi|^N$ appearing in the Green's functions Γ_{0N}, are small parameters for large N. Henceforth, we will refer to such quantities as "exponentially small," implying exponential decay with chain length (or number of sites).

Expanding Equation (49) in terms of ω and $\chi_\pm^N$, which we treat as small, as explained above, we obtain

$$0 = \det \left[-\hat{\mathrm{I}}\omega \frac{t}{|\Delta|(4t^2 - \mu^2)} - \right.$$
$$- \begin{pmatrix} 1 & \frac{\Delta}{|\Delta|} \\ \frac{\Delta^*}{|\Delta|} & 1 \end{pmatrix} \frac{1}{\omega} \frac{|\Delta|(4t^2 - \mu^2)}{2t(4(t^2 - |\Delta|^2) - \mu^2)} \times \quad (53)$$
$$\left. \times \left(\chi_+^{N+1} - \chi_-^{N+1} \right)^2 \right],$$

where $\hat{\mathrm{I}}$ is the identity matrix. The solution $\omega = 0$ corresponds to the pole of the Green's function, which exists only on the semi-infinite chain segments. The other pair of solutions has finite but small energies $\omega = \pm\omega_0$, where

$$\omega_0 = \frac{|\Delta|(4t^2 - \mu^2)}{it\sqrt{4(t^2 - |\Delta|^2) - \mu^2}} \times$$
$$\times \left(\chi_+^{N+1} - \chi_-^{N+1} \right). \quad (54)$$

Here we see that this solution satisfies the approximations we made, if $\left|\chi_\pm^{N+1}\right| \ll 1$. Considering Equation (51), the condition $t^2 = ((\mu/2)^2 + \Delta^2)$ separates the two regions with oscillating and non-oscillating solutions for ω_0. If ω_0 crosses zero with varying μ, this implies a change in fermion parity, as discussed in [36].

The leading term in the expansion of the Green's function $\Gamma_{nm}^R(\omega)$ near $\omega \to \pm\omega_0$, which in quantum mechanics would describe the spatial structure of the wavefunctions of the two localized states, takes the following form:

$$\Gamma_{nm}^R(\omega) = -\frac{\omega}{(\omega + i\delta)^2 - (\omega_0)^2} \frac{|\Delta|(4t^2 - \mu^2)}{2t(4(t^2 - |\Delta|^2) - \mu^2)} \times$$
$$\times \left((\chi_+^n - \chi_-^n)(\chi_+^m - \chi_-^m) \begin{pmatrix} 1 & \frac{\Delta}{|\Delta|} \\ \frac{\Delta^*}{|\Delta|} & 1 \end{pmatrix} + (\chi_+^{N+1-n} - \chi_-^{N+1-n})(\chi_+^{N+1-m} - \chi_-^{N+1-m}) \begin{pmatrix} 1 & -\frac{\Delta}{|\Delta|} \\ -\frac{\Delta^*}{|\Delta|} & 1 \end{pmatrix} \right) -$$
$$- \frac{\omega_0}{(\omega + i\delta)^2 - (\omega_0)^2} \frac{|\Delta|(4t^2 - \mu^2)}{2t(4(t^2 - |\Delta|^2) - \mu^2)} \times \quad (55)$$
$$\times \left((\chi_+^n - \chi_-^n)(\chi_+^{N+1-m} - \chi_-^{N+1-m}) \begin{pmatrix} 1 & -\frac{\Delta}{|\Delta|} \\ \frac{\Delta^*}{|\Delta|} & -1 \end{pmatrix} + (\chi_+^{N+1-n} - \chi_-^{N+1-n})(\chi_+^m - \chi_-^m) \begin{pmatrix} 1 & \frac{\Delta}{|\Delta|} \\ -\frac{\Delta^*}{|\Delta|} & -1 \end{pmatrix} \right).$$

Diagonal elements Γ_{nn} show the spatial distribution of density in localized states. In the limit $\Delta = t$, only Γ_{11} and Γ_{NN} remain non-zero, since Equation (52) gives

$$\chi_+^n \propto \chi_-^n \propto |t - |\Delta||^{n/2}.$$

In the high-symmetry case $\mu = 0$ and $|\Delta| \to t$, the energy levels are equal:

$$\omega_0 = \frac{4|\Delta|t}{t + |\Delta|} \left(\frac{t - |\Delta|}{t + |\Delta|} \right)^{\frac{N}{2}} \times$$
$$\times \sin\left(\frac{\pi(N+1)}{2} \right) \to 0. \quad (56)$$

As noted earlier (see, for example, [37]), for an odd number of sites is equal to zero for any values of t and Δ.

REFERENCES

1. *A. Yu. Kitaev*. PhysicsUsp. **44**, 131, (2001)
2. *A. F. Andreev*, Sov. Phys. JETP **19**, 1228 (1964); 49, 655 (1966); 22, 455 (1966)
3. *V. V. Val'kov, M. S. Shustin, S. V. Aksenov et al.*, Phys. Usp. **65,** 2 (2022)
4. *V. Mourik, K. Zuo, S. M. Frolov et al.* Science **336**, 1003 (2012).
5. *M. T. Deng et al.* Science **354**, 1557 (2016).
6. *H. J. Suominen, M. Kjaergaard, A. R. Hamilton et al.* Phys. Rev. Lett. **119**, 176805 (2017)
7. *E. Prada, P. San-Jose, M. W. A. de Moor et al*. Nature Reviews Physics **2**, 575 (2020)
8. *D. Aasen, M. Hell, R. V. Mishmash et al*. Physical Review X **6**, 031016 (2016)
9. *T. Karzig, Y. Oreg, G. Refael and M. H. Freedman*, Physical Review X **6**, 031019 (2016)
10. *F. Pientka, L. Jiang, D. Pekker et al*. New Journal of Physics **15**, 115001 (2013)
11. *T. Hyart, B. van Heck, I. C. Fulga et al.* Phys. Rev. B **88**, 035121 (2013)
12. *L. Jiang, D. Pekker, J. Alicea et al.* Phys. Rev. Lett. **107**, 236401 (2011)
13. *B. Van Heck, A. R. Akhmerov, F. Hassler et al*. New Journal of Physics **14**, 035019 (2012)
14. *M. Valentini, M. Borovkov, E. Prada et al.* Nature **612**, 442 (2022)
15. *S. M. Albrecht, A. P. Higginbotham, M. Madsen et al.* Nature **531**, 206 (2016)
16. *S. D. Sarma, M. Freedman, and C. Nayak*, npj Quantum Inf. **1**, 15001 (2015)
17. *C. Nayak, S. H. Simon, A. Stern et al.* Rev. Mod. Phys. **80** 1083 (2008)
18. *J. Jin, and X. Q. Li*, New J. Phys. **24** 093009 (2022)
19. *L. Arrachea, G. S. Lozano, and A. A. Aligia*, Phys. Rev. B **80**, 014425 (2009).
20. *A. Zazunov, R. Egger, and A. Levy*, Yeyati Phys. Rev. B **94**, 014502 (2016).
21. *A. I. Rusinov*, Sov. Phys. JETP. **29**, 1101 (1969)
22. *P. I. Arseyev, and B. A. Volkov*, Sol.St.Comm. **78**, 373 (1991)
23. *C. Caroli, R. Combescot, P. Nozieres, and D. SaintJames*, J. Phys. C: Solid State Phys. **4**, 916 (1971)
24. *K. Flensberg*, Phys. Rev. B **82**, 180516 (2010)
25. *S. Smirnov*, Phys. Rev. B **105**, 205430 (2022)
26. *L. Fu*, Phys. Rev. Lett. **104**, 056402 (2010)
27. *Y. Meir, and N. S. Wingreen*, Phys. Rev. Lett. **68**, 2512 (1992)
28. *N. Leumer, M. Grifoni, B. Muralidharan, and M. Marganska*, Physical Review B. **103**, 165432 (2021)
29. *C. Caroli, R. Combescot, P. Nozieres, and D. SaintJames*, J. Phys. C: Solid State Phys. **5**, 21 (1972)
30. *P. I. Arseev, N. S. Maslova*, Physics-Usp. **53**, 1151 (2010)
31. *S. N. Thomas, S. D. Sarma, and J. D. Sau*, Phys. Rev. B **106**, 174501 (2022)
32. *Y. H. Lai, S. D. Sarma, and J. D. Sau*, Phys. Rev. B **106**, 094504 (2022)
33. *S. Hegde et al.*, New J. Phys. **17**, 053036 (2015)
34. *K. Kawabata, R. Kobayashi, N. Wu, and H. Katsura*, Phys. Rev. B **95**, 195140 (2017)

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 124–135

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

SUPRATHERMAL ELECTRON TRANSPORT IN WEAKLY AND STRONGLY MAGNETIZED ASTROPHYSICAL PLASMAS INCLUDING COULOMB COLLISIONAL EFFECTS

© 2025 Ji-Hoon Ha

Korea Space Weather Center, Korea AeroSpace Administration, Jeju, South Korea
e-mail: hjhspace223@gmail.com

Received July 01, 2024

Revised August 15, 2024

Accepted August 20, 2024

Abstract. This study examines electron transport in astrophysical plasmas mediated by Coulomb collisions and collisionless wave-particle interactions, using a kinetic transport model that incorporates spectral evolutions through these interactions. It investigates the transport of suprathermal electrons via whistler turbulence and the effects of plasma magnetization. Key findings indicate that in strongly magnetized plasmas, diffusion timescales in pitch angle space become saturated at large pitch angles, independent of increasing magnetic field strength. Conversely, in weakly magnetized plasmas, these timescales decrease with decreasing magnetic field strength, enhancing electron transport in velocity space. The study also identifies minimum conditions for resonant scattering, dominated by wave-particle interactions over Coulomb collisions, which depend on Coulomb collision effects and the power-law slope of the whistler turbulence spectrum. These findings have applications in weakly magnetized astrophysical plasmas, from the relatively strong magnetic fields of the interplanetary medium to the very weak magnetic fields of the intracluster medium.

DOI: 10.31857/S0044451025011le6

1. INTRODUCTION

Plasma physics is essential for understanding various astrophysical and laboratory phenomena, where electron transport significantly influences the behavior and evolution of plasma systems. In the field of fusion plasma, plasma heating and current drive have been primarily examined to maintain the conditions necessary for the magnetic confinement of plasmas [1]. It has been demonstrated that the propagation and damping of radiofrequency waves, including ion cyclotron, electron cyclotron, and lower-hybrid waves, produce energetic ions and electrons through Landau and cyclotron damping, which leads to current drive generation in the plasma system. Along with such collisionless damping, the collisional relaxation of energetic particles is involved in the evolution of particle distribution in the plasma system. Likewise, collisionless wave-particle interactions and collisional relaxation also play a crucial role in particle transport in astrophysical plasmas. Indeed, turbulence and the associated plasma instabilities are ubiquitous in astrophysical plasmas, and understanding energy transport through such turbulence is a long-standing problem [2–5].

Plasma phenomena and their dynamical evolution in space and astrophysical plasmas depend on the magnetization, defined as follows:

$$\frac{\omega_{pe}}{\Omega_e} = \frac{\sqrt{4\pi n_0 e^2 / m_e}}{eB_0 / m_e c} \propto \frac{\sqrt{n_0}}{B_0}, \qquad (1)$$

where

$$\omega_{pe} = \sqrt{4\pi n_0 e^2 / m_e}, \quad \Omega_e = eB_0 / m_e c$$

stand for the plasma frequency and electron gyrofrequency, respectively, and these quantities depend on the plasma density n_0 and magnetic field B_0. Thus, the phenomena associated with plasma physics have been examined across a wide range of magnetization factors [6–11]. For instance, the characteristics of plasma instabilities in space plasma depend on the properties of the medium, such as strongly magnetized plasma in the solar atmosphere near the Sun ($\omega_{pe} / \Omega_e < 1$) and weakly magnetized plasmas in the solar wind propagating

toward Earth ($\omega_{pe}/\Omega_e > 1$) [6, 7]. Additionally, a wide range of ω_{pe}/Ω_e can be adopted to model the pulsar wind propagation from the strongly magnetized magnetosphere of a pulsar to the weakly magnetized pulsar wind nebulae propagating toward the interstellar medium [8, 9]. Furthermore, rigorous theories have been proposed for kinetic turbulence and their roles in particle heating through energy transfer in ambient astrophysical environments, including weakly magnetized media such as interplanetary, interstellar, and intracluster media ($\omega_{pe}/\Omega_e \gg 1$) [10, 11].

Understanding turbulence and dynamical evolution in various astrophysical media is crucial for comprehending particle transport across strongly magnetized to weakly magnetized plasmas, which is essential to examine the nature of plasma distribution in various space and astrophysical plasmas. The mechanisms behind particle transport in space weather have been particularly examined so far. Indeed, suprathermal electrons have been observed by the Parker Solar Probe in the interplanetary medium; these electrons are expected to originate in the solar corona and escape into the interplanetary medium along open magnetic field lines [12, 13]. While particle transport in plasmas has primarily been attributed to Coulomb collisions, observational evidence of suprathermal electrons highlights the importance of collisionless wave-particle interactions. In this regard, recent theoretical studies have proposed a kinetic model based on the Fokker-Planck equation, including wave-particle interactions mediated by plasma turbulence [14–23]. For instance, Kim et al. [14] highlighted that the persistence of a non-Maxwellian distribution in the solar wind could be exhibited through wave-particle interactions due to Langmuir turbulence in the absence of Coulomb collisions (see also [15]). Tang et al. [16] incorporated Coulomb collisional effects along with wave-particle interaction terms into the kinetic model and showed that Coulomb collisions predominantly transport core electrons following a Maxwellian distribution, whereas suprathermal electrons are preferentially accelerated through whistler turbulence. Simulation studies using the particle-in-cell (PIC) method have also shown the formation of suprathermal electrons through whistler turbulence [24, 25]. These findings are consistent with observational evidence of suprathermal electrons in interplanetary space [12, 13].

Despite the considerable progress mentioned above, several gaps persist in our understanding, particularly regarding how these mechanisms operate under different plasma magnetization conditions. Notably, the plasma parameters, including magnetization, differ between interplanetary space and other astrophysical media such as interstellar and intracluster media. Consequently, plasma phenomena related to particle transport could also differ. While simulation studies using kinetic plasma simulations have demonstrated possible acceleration mechanisms through collisionless shocks and turbulence in various astrophysical media [26–31], it is essential to understand the transport of such accelerated particles in these media to demonstrate the persistence of non-Maxwellian distributions.

In this context, this work aims to improve our understanding of particle transport theory based on the kinetic transport equation and whistler turbulence under different plasma magnetization conditions relevant to various astrophysical media. To achieve this, we adopt a kinetic transport model that incorporates the spectral evolution influenced by both Coulomb collisions and wave-particle interactions, as proposed in previous works [16–19]. By examining how suprathermal electrons are transported through whistler turbulence under varying degrees of plasma magnetization, we extend the applicability of the kinetic transport model to various astrophysical environments. This work reveals distinct behaviors in diffusion timescales for weakly and strongly magnetized plasmas, with significant implications for electron transport dynamics. Additionally, we identify minimum conditions for resonant scattering dominated by wave-particle interactions over Coulomb collisions, highlighting dependencies on Coulomb collision effects and the power-law slope of the whistler turbulence spectrum. This comprehensive approach allows us to explore diffusion timescales in both velocity and pitch angle space, providing new insights into the underlying processes governing electron transport in plasmas.

2. DESCRIPTION OF THE KINETIC MODEL

The evolution of the electron velocity distribution function in astrophysical environments has been examined using the kinetic transport equation [16–19]. The electromagnetic interaction in a typical astrophysical environment includes the electric force and the Lorentz force, which are described as follows:

$$\mathbf{a} = -\frac{e\mathbf{E}}{m_e} - \frac{e}{m_e}(\mathbf{v}\times\mathbf{B}) = \mathbf{a}_r(\mathbf{r}) + \mathbf{a}_L. \qquad (2)$$

Here, e and m_e are the electric charge and the mass of electrons, and $\mathbf{E}$ and $\mathbf{B}$ denote the electric and magnetic fields, respectively. $\mathbf{a}_r(\mathbf{r})$ is the radial component of the acceleration due to the electric force, whereas $\mathbf{a}_L$ is the non-radial component due to the Lorentz force. Using the acceleration $\mathbf{a}$ due to the external forces along with the terms responsible for Coulomb collisions and wave-particle interactions of kinetic turbulence, the kinetic transport equation can be described as follows:

$$\frac{\partial f(\mathbf{r},\mathbf{v},t)}{\partial t} + (\mathbf{v}\cdot\nabla_{\mathbf{r}})f(\mathbf{r},\mathbf{v},t) + (\mathbf{a}\cdot\nabla_{\mathbf{v}})f(\mathbf{r},\mathbf{v},t) = \left(\frac{\delta f}{\delta t}\right)_{cc} + \left(\frac{\delta f}{\delta t}\right)_{wp}. \qquad (3)$$

Here, the electron velocity distribution function is expressed in the position ($\mathbf{r}$), velocity ($\mathbf{v}$) and time (t) domains, and $(\delta f/\delta t)_{cc}$ and $(\delta f/\delta t)_{wp}$ include the effects of Coulomb collisions and kinetic turbulence, respectively. In the coordinates of the radial distance r, the velocity v, and the parameter including the pitch angle θ between the velocity and magnetic field vectors ($\mu \equiv \cos\theta$), Equation (3) becomes

$$\frac{\partial f}{\partial t} + v\propto\frac{\partial f}{\partial r} + a_r(r)\left(\propto\frac{\partial f}{\partial v} + \frac{(1-\propto^2)}{v}\frac{\partial f}{\partial\propto}\right) + \frac{v}{r}(1-\mu^2)\frac{\partial f}{\partial\mu} = \left(\frac{\delta f}{\delta t}\right)_{cc} + \left(\frac{\delta f}{\delta t}\right)_{wp}. \qquad (4)$$

The Coulomb collisions with Maxwellian backgrounds of electrons and protons have been employed in the solar wind environments [16]. The term associated with the Coulomb collisions [32] can be expressed as:

$$\left(\frac{\delta f}{\delta t}\right)_{cc} = c_{v,e}\left\{\left[\mathrm{erf}\left(\frac{v}{v_{\mathrm{th},e}}\right) - G\left(\frac{v}{v_{\mathrm{th},e}}\right)\right]\times \frac{1}{2v^3}\frac{\partial}{\partial\propto}\left[(1-\propto^2)\frac{\partial f}{\partial\propto}\right] + \frac{1}{v^2}\frac{\partial}{\partial v}\left[G\left(\frac{v}{v_{\mathrm{th},e}}\right)v\frac{\partial f}{\partial v}\right] + \frac{1}{v^2}\frac{\partial}{\partial v}\left[\frac{2v^2}{v_{\mathrm{th},e}^2}G\left(\frac{v}{v_{\mathrm{th},e}}\right)f\right]\right\} + c_{v,p}\left\{\left[\mathrm{erf}\left(\frac{v}{v_{\mathrm{th},p}}\right) - G\left(\frac{v}{v_{\mathrm{th},p}}\right)\right]\times \frac{1}{2v^3}\frac{\partial}{\partial\propto}\left[(1-\propto^2)\frac{\partial f}{\partial\propto}\right] + \frac{1}{v^2}\frac{\partial}{\partial v}\left[G\left(\frac{v}{v_{\mathrm{th},p}}\right)v\frac{\partial f}{\partial v}\right] + \frac{1}{v^2}\frac{\partial}{\partial v}\left[\frac{2v^2}{v_{\mathrm{th},p}^2}\frac{m_e}{m_p}G\left(\frac{v}{v_{\mathrm{th},e}}\right)f\right]\right\}, \qquad (5)$$

where m_p/m_e is the proton-to-electron mass ratio and $v_{\mathrm{th},e}$ and $v_{\mathrm{th},p}$ are the thermal velocities of the background Maxwellian electrons and protons. $\mathrm{erf}(x)$ and $G(x)$ are the error function and the Chandrasekhar function, respectively. The collision frequencies corresponding to the collisions with the Maxwellian background electrons ($c_{v,e}$) and protons ($c_{v,p}$) are given by:

$$c_{v,e} = \frac{4\pi n_0 e^4 \ln\Lambda}{m_e^2}, \qquad (6)$$

$$c_{v,p} = \frac{4\pi n_0 e^4 \ln\Lambda}{m_p^2}, \qquad (7)$$

where n_0 and $\ln\Lambda$ are the plasma density and the Coulomb logarithm.

To model the terms for wave-particle interaction, we consider the resonant scattering of electrons by right-handed polarized whistler waves as a main wave-particle interaction mechanism in the turbulent plasma system. Considering the cyclotron resonance of electrons with waves propagating parallel to the guiding magnetic field $\mathbf{B}_0$, the resonant particles satisfy the following condition:

$$\omega_r(k) = v\mu k_{\parallel} + n\Omega_e, \qquad (8)$$

where ω_r and k are the oscillatory wave frequency and the wavenumber, respectively, and $\Omega_e = |e|B_0/m_e c$ is the electron gyrofrequency. The integer $n \neq 0$ must be finite for cyclotron resonance through the parallel waves. In the whistler regime ($\omega_r < \Omega_e$), the magnetic power spectrum [18, 22] can be described as follows:

$$P_B(k) = A\frac{c}{\Omega_e}\left|\frac{kc}{\Omega_e}\right|^{-s}, \qquad (9)$$

where A is the normalization constant, and the spectral index s is expected not to exceed 2 [22]. The evolution of the electron distribution function due to wave-particle interaction through whistler turbulence [16–19] can be expressed as

$$\left(\frac{\delta f}{\delta t}\right)_{wp} = \frac{\partial}{\partial \mu}\left(D_{\mu\mu}\frac{\partial f}{\partial \mu} + \frac{1}{m_e}D_{\mu v}\frac{\partial f}{\partial v}\right) + \\ + \frac{1}{v^2}\frac{\partial}{\partial v}\left(v^2\left(\frac{1}{m_e}D_{\propto v}\frac{\partial f}{\partial \propto} + \frac{1}{m_e^2}D_{vv}\frac{\partial f}{\partial v}\right)\right). \quad (10)$$

The diffusion tensor for nonrelativistic electrons is expressed as:

$$\bar{D}_{vv} \equiv \frac{D_{vv}}{\Omega_e (m_e c)^2} = \frac{\pi}{3}\frac{A}{a}\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-1}{3}}(1-\mu^2), \quad (11)$$

$$\bar{D}_{\mu v} \equiv \frac{D_{\mu v}}{\Omega_e (m_e c)} = \frac{D_{v\mu}}{\Omega_e (m_e c)} = \\ = -\frac{\pi}{3}\frac{A}{a}\left[\frac{\mu}{|\mu|}\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-2}{3}} + \frac{\mu}{\beta}\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-1}{3}}\right](1-\mu^2), \quad (12)$$

$$\bar{D}_{\mu\mu} \equiv \frac{D_{\mu\mu}}{\Omega_e} = \frac{\pi}{3}\frac{A}{a}\times \\ \times\left[\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-3}{3}} + 2\frac{\mu}{|\mu|}\frac{\mu}{\beta}\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-2}{3}} + \right. \\ \left. + \left(\frac{\mu}{\beta}\right)^2\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-1}{3}}\right](1-\mu^2). \quad (13)$$

Here, we used dimensionless parameters, $\beta = v/c$ and $a = \omega_{pe}^2/\Omega_e^2$ with the plasma frequency $\omega_{pe} = \sqrt{4\pi n_0 e^2/m_e}$. To consider both weakly magnetized plasmas such as interplanetary, interstellar, and intracluster media ($a \gg 1$) and strongly magnetized plasmas near the stellar magnetosphere ($a < 1$), we examine the properties of wave-particle interactions mediated by whistler turbulence over a wide range of parameter a.

In the kinetic model described by Equation (4), the detailed evolution mediated by Coulomb collisions and wave-particle interactions depends on the initial electron distribution. The electron distribution of thermal plasma is typically modeled as Maxwellian, given by:

$$f_{\mathrm{th},e}(v) = \frac{n_0}{\neq^{3/2} v_{\mathrm{th},e}^3}\exp\left[-\left(\frac{v}{v_{\mathrm{th},e}}\right)^2\right]. \quad (14)$$

While the Maxwellian distribution is suitable for describing the medium in the absence of nonlinear processes such as plasma and magnetohydrodynamic (MHD) waves, shocks, and turbulence, it has been demonstrated that plasma processes associated with such phenomena can accelerate particles. This particle energization results in a distribution that deviates from Maxwellian, known as the kappa distribution [33–35]. The electron kappa distribution is defined as:

$$f_{\kappa,e}(v) = \frac{n_0}{\pi^{3/2} v_{\mathrm{th},e}^3}\frac{\Gamma(\kappa+1)}{(\kappa-3/2)^{3/2}\Gamma(\kappa-1/2)}\times \\ \times\left[1+\frac{1}{(\kappa-3/2)}\left(\frac{v}{v_{\mathrm{th},e}}\right)^2\right]^{-(\kappa+1)}, \quad (15)$$

where $\Gamma(x)$ is the Gamma function and the parameter κ determines the slope of the suprathermal distribution. For $v \gg v_{\mathrm{th},e}$, the kappa distribution follows a power-law form,

$$f_{\kappa,e}(v) \propto v^{-2(\kappa+1)}.$$

A smaller value of κ results in a flatter particle distribution, whereas a larger value of κ makes the kappa distribution closer to Maxwellian. In the subsequent section, we explore how the initial slope of the electron distribution function influences electron transport through whistler turbulence, taking into account the dependence on magnetization.

It is noteworthy that the nature of plasma turbulence and wave-particle interaction mediated by such turbulence could be substantially different from the interpretation obtained through linear theory [36, 37]. Specifically, the effects of nonlinear processes on energy dissipation by whistler waves have been examined through PIC simulations [38, 39]. According to the results of these numerical simulations, the significance of nonlinear damping of whistler waves depends on the fluctuation energy of the turbulence and the magnetization of the plasma system [38]. In weakly magnetized plasma, linear damping dominates over nonlinear damping, indicating that the theory developed in the linear regime could be applicable for examining wave-particle interaction through whistler turbulence. In strongly magnetized plasma, when the turbulent

fluctuation (δB) is sufficiently weak (i.e., $\delta B \leq B_0$), linear theory could be applicable. In this regard, the kinetic model in this work could be suitable for weak turbulence systems in space and astrophysical environments. For systems with strong turbulence ($\delta B \geq B_0$), nonlinear processes should be taken into account in the model, which is beyond the scope of this paper.

3. ELECTRON TRANSPORT THROUGH WAVE-PARTICLE INTERACTION AND ITS DEPENDENCE ON THE MAGNETIZATION OF THE PLASMA SYSTEM

Comparison of $\tau_{\mu v} / \tau_{vv}$ (upper panels) and $\tau_{\mu\mu} / \tau_{vv}$ (lower panels) across parameter space. The plots depict variations with respect to electron velocity β ranging from 10^{-3} to 10^{-1}, and magnetization parameter a spanning from 10^{-4} to 10^{4}. Larger values of a indicate weakly magnetized plasmas, whereas smaller values denote strongly magnetized plasma

Firstly, we examine the acceleration timescales through whistler turbulence and their dependence on the magnetic field strength using the three diffusion coefficients. The acceleration timescales can be derived as follows:

$$\frac{\tau_{vv}}{\Omega_e^{-1}} \equiv \frac{\gamma_e^2 m_e^2 v^2}{\Omega_e^{-1} D_{vv}} = \frac{3a}{A\pi}\beta^2 \left(\frac{\beta|\mu|}{a}\right)^{\frac{1-s}{3}} (1-\mu^2)^{-1}, \quad (16)$$

$$\frac{\tau_{\mu v}}{\Omega_e^{-1}} \equiv \left|\frac{\gamma_e m_e v}{\Omega_e^{-1} D_{\mu v}}\right| =$$

$$= \frac{3a}{A\pi}\beta \left|\frac{\mu}{|\mu|}\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-2}{3}} + \frac{\mu}{\beta}\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-1}{3}}\right|^{-1} \times$$

$$\times (1-\propto^2)^{-1}, \quad (17)$$

$$\frac{\tau_{\mu\mu}}{\Omega_e^{-1}} \equiv \left|\frac{1}{\Omega_e^{-1} D_{\mu\mu}}\right| =$$

$$= \frac{3a}{A\pi}\left|\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-3}{3}} + \frac{2\mu}{|\mu|}\frac{\mu}{\beta}\left(\frac{\beta|\mu|}{a}\right)^{\frac{s-2}{3}} + \right.$$

$$\left. + \left(\frac{\mu}{\beta}\right)^2 \left(\frac{\beta|\mu|}{a}\right)^{\frac{s-1}{3}}\right|^{-1} (1-\mu^2)^{-1}, \quad (18)$$

where γ_e is the Lorentz factor, which is approximately 1 for nonrelativistic particles. To assess the relative importance of pitch angle scattering, the following ratios were calculated:

$$\frac{\tau_{\mu v}}{\tau_{vv}} = \left|\frac{\beta\mu}{|\mu|}\left(\frac{\beta|\mu|}{a}\right)^{-\frac{1}{3}} + \mu\right|^{-1}, \quad (19)$$

$$\frac{\tau_{\mu\mu}}{\tau_{vv}} = \left|\beta^2\left(\frac{\beta|\mu|}{a}\right)^{-\frac{2}{3}} + \frac{2\beta\mu^2}{|\mu|}\left(\frac{\beta|\mu|}{a}\right)^{-\frac{1}{3}} + \mu^2\right|^{-1}. \quad (20)$$

In a strongly magnetized plasma ($a \to 0$), the ratios simplify to:

$$\frac{\tau_{\mu v}}{\tau_{vv}} \to |\mu|^{-1}, \frac{\tau_{\mu\mu}}{\tau_{vv}} \to |\mu|^{-2}, \quad (21)$$

indicating that the relative importance of diffusion in pitch angle space is independent of the particle velocity β and magnetic field strength parametrized by a once the particles satisfy the resonant condition. Given that the pitch angle parameter satisfies $|\propto| < 1$, the following relations hold true in strongly magnetized plasmas:

$$\tau_{vv} < \tau_{\mu v} < \tau_{\mu\mu}. \quad (22)$$

In weakly magnetized plasmas ($a \gg 1$), however, the ratios of these characteristic timescales may vary depending on the particle velocity β and magnetic field strength a.

Fig. 1 shows $\tau_{\mu v} / \tau_{vv}$ and $\tau_{\mu\mu} / \tau_{vv}$ as functions of electron velocity β and magnetization a. A few points were noted: (1) In weakly magnetized plasmas ($a \gg 1$), diffusion processes in the pitch angle space become prominent, whereas a saturated behavior is observed for particle acceleration in sufficiently strong magnetic fields ($a \ll 1$). (2) The dependence on magnetic field strength is more pronounced for accelerating electrons with higher β. Particularly, panels (a) and (d) show that $\tau_{\mu v} / \tau_{vv}$ and $\tau_{\mu\mu} / \tau_{vv}$ exhibit similar asymptotic behaviors for small β and large pitch angles $|\mu| > 0.5$, irrespective of A. Conversely, panels (b), (c), (e), and (f) illustrate that the effects of magnetic field strength on pitch angle scattering are more significant for electrons with larger β. (3) In strongly magnetized plasmas ($a < 1$), $\tau_{\mu v}$ and $\tau_{\mu\mu}$ increase as the pitch angle $|\mu|$ decreases, whereas the opposite behavior is observed

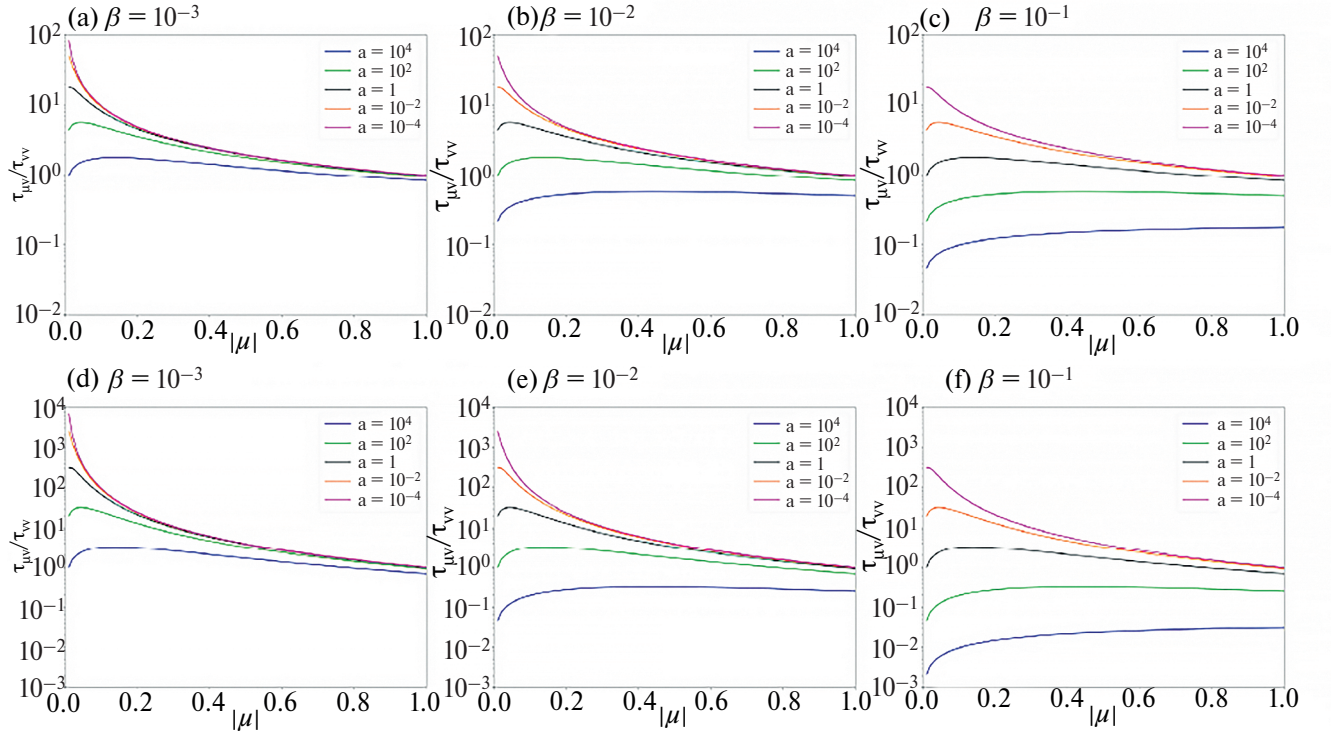

Fig. 1. Comparison of $\tau_{\mu v}/\tau_{vv}$ (upper panels) and $\tau_{\mu\mu}/\tau_{vv}$ (lower panels) across parameter space. The plots depict variations with respect to electron velocity β ranging from 10^{-3} to 10^{-1}, and magnetization parameter a spanning from 10^{-4} to 10^{4}. Larger values of a indicate weakly magnetized plasmas, whereas smaller values denote strongly magnetized plasma

in weakly magnetized plasmas ($a > 1$). This indicates that wave-particle interactions are influenced by the magnetic field strength of the background medium.

Next, we examine the conditions under which the acceleration timescales are dominated by wave-particle interactions over Coulomb collisions. Assuming fixed background temperatures (constant $v_{\mathrm{th},e}$ and $v_{\mathrm{th},p}$), these regimes depend on the magnetic field strength and the initial distribution of suprathermal electrons. Considering the diagonal terms in $\left(\delta f / \delta t\right)_{cc}$ and $\left(\delta f / \delta t\right)_{wp}$ for velocity space diffusion, we have the following expressions:

$$\left(\frac{\delta f}{\delta t}\right)_{cc} \approx c_{v,e}\left\{\frac{1}{v^2}\frac{\partial}{\partial v}\left[G\left(\frac{v}{v_{\mathrm{th},e}}\right)v\frac{\partial f}{\partial v}\right]+\right.$$

$$\left.+\frac{1}{v^2}\frac{\partial}{\partial v}\left[\frac{2v^2}{v_{\mathrm{th},e}^2}G\left(\frac{v}{v_{\mathrm{th},e}}\right)f\right]\right\}+$$

$$+c_{v,p}\left\{\frac{1}{v^2}\frac{\partial}{\partial v}\left[G\left(\frac{v}{v_{\mathrm{th},p}}\right)v\frac{\partial f}{\partial v}\right]+\right.$$

$$\left.=\frac{1}{v^2}\frac{\partial}{\partial v}\left[\frac{2v^2}{v_{\mathrm{th},p}^2}\frac{m_e}{m_p}G\left(\frac{v}{v_{\mathrm{th},e}}\right)f\right]\right\}, \qquad (23)$$

$$\left(\frac{\delta f}{\delta t}\right)_{wp} \approx \frac{1}{v^2}\frac{\partial}{\partial v}\left[v^2\left(\frac{1}{m_e^2}D_{vv}\frac{\partial f}{\partial v}\right)\right]=$$

$$=\frac{1}{v^2}\left[\frac{2v}{m_e^2}D_{vv}\frac{\partial f}{\partial v}+\frac{\partial D_{vv}}{\partial v}\frac{v^2}{m_e^2}\frac{\partial f}{\partial v}+\right.$$

$$\left.+\frac{v^2}{m_e^2}D_{vv}\frac{\partial^2 f}{\partial v^2}\right]. \qquad (24)$$

For $v \gg v_{\mathrm{th},e}$, the Chandrasekhar function can be approximated as $G(v / v_{\mathrm{th},e}) \approx (v / v_{\mathrm{th},e})^{-2} / 2$ and Equation (23) simplifies to:

$$\left(\frac{\delta f}{\delta t}\right)_{cc} \approx c_{v,e}\left\{\frac{1}{v^2}\frac{\partial f}{\partial v}+\right.$$

$$\left.+\frac{1}{2}\left(\frac{v_{\mathrm{th},e}}{v}\right)^2\left[-\frac{1}{v^2}\frac{\partial f}{\partial v}+\frac{1}{v}\frac{\partial^2 f}{\partial v^2}\right]\right\}+$$

$$+c_{v,p}\left\{\frac{1}{v^2}\frac{m_e}{m_p}\frac{\partial f}{\partial v}+\right.$$

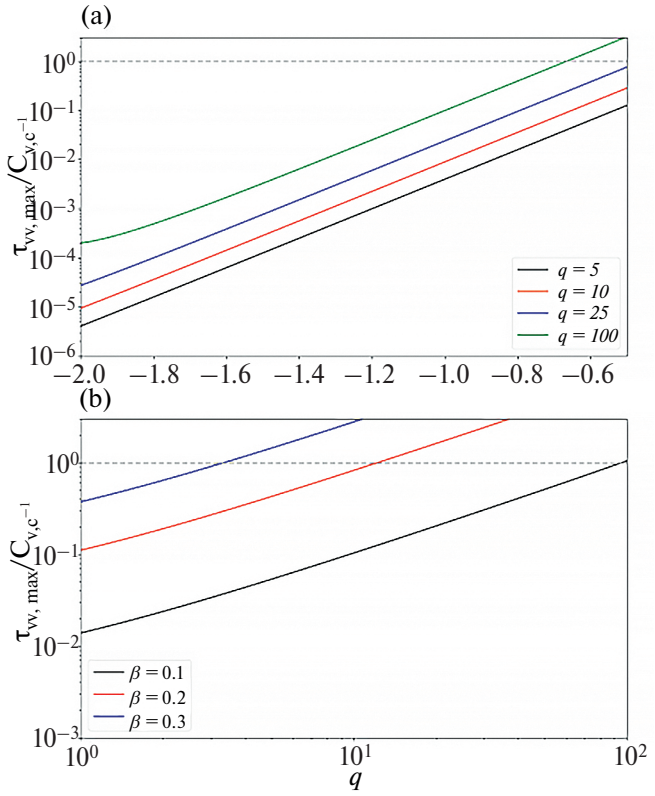

Fig. 2. *a* – Maximum acceleration timescale, $\tau_{\mu v,\max}$, plotted against β for four different q values. **b** – $\tau_{\mu v,\max}$ shown for three different β values across various q values. The electron thermal velocity is set as $v_{\mathrm{th},e}/c = 10^{-3}$. Gray lines indicate $\tau_{\mu v,\max}/c_{v,e}^{-1} = 1$

$$+\frac{1}{2}\left(\frac{v_{\mathrm{th},p}}{v}\right)^2\left[-\frac{1}{v^2}\frac{\partial f}{\partial v}+\frac{1}{v}\frac{\partial^2 f}{\partial v^2}\right]\Bigg\}. \qquad (25)$$

Assuming the suprathermal electrons follow a kappa distribution function, the distribution of high-energy electrons with $v \gg v_{\mathrm{th},e}$ approximates to a power-law tail, $f \propto v^{-q}$. The derivatives of f are expressed as follows:

$$\frac{\partial f}{\partial v} = -qv^{-1}f, \qquad (26)$$

$$\frac{\partial^2 f}{\partial v^2} = q(q+1)v^{-2}f. \qquad (27)$$

Using Equations (26) and (27), Equations (24) and (25) can be rewritten as

$$\left(\frac{\delta f}{\delta t}\right)_{cc} \approx c_{v,e}\left\{-\frac{qf}{v^3}+\frac{1}{2}\left(\frac{v_{\mathrm{th},e}}{v}\right)^2\left[\frac{q(q+2)f}{v^3}\right]\right\}+$$
$$+c_{v,p}\left\{-\frac{m_e}{m_p}\frac{qf}{v^3}+\frac{1}{2}\left(\frac{v_{\mathrm{th},p}}{v}\right)^2\left[\frac{q(q+2)f}{v^3}\right]\right\}, \qquad (28)$$

$$\left(\frac{\delta f}{\delta t}\right)_{wp} \approx \frac{1}{v^2}\left[\frac{q(q-1)D_{vv}}{m_e^2}f - \frac{qvD_{vv}(s-1)\beta^{(s-4)/3}}{3m_e^2 c}f\right]. \qquad (29)$$

Electrons gain energy when

$$\left(\delta f/\delta t\right)_{cc} + \left(\delta f/\delta t\right)_{wp} \geq 0.$$

In this case, we obtain the following inequality for D_{vv}:

$$D_{vv} \geq \left[\frac{q(q-1)}{m_e^2 v^2} - \frac{q(s-1)\beta^{(s-4)/3}}{3m_e^2 vc}\right]^{-1} \times$$
$$\times\left\{c_{v,e}\left[\frac{q}{v^3}-\frac{1}{2}\left(\frac{v_{\mathrm{th},e}}{v}\right)^2\left(\frac{q(q+2)}{v^3}\right)\right]+\right.$$
$$\left.+c_{v,p}\left[\frac{m_e}{m_p}\frac{q}{v^3}-\frac{1}{2}\left(\frac{v_{\mathrm{th},p}}{v}\right)^2\left(\frac{q(q+2)}{v^3}\right)\right]\right\}. \qquad (30)$$

Using the inequality (30), we examine how the slope of the initial distribution of suprathermal electrons could influence the relative importance between Coulomb collisions and wave-particle interactions. For nonrelativistic electrons where $v_{\mathrm{th},e}/c \ll \beta \ll 1$ (or the Lorentz factor $\gamma_e \approx 1$), the acceleration timescale (τ_{vv}) satisfies

$$\tau_{vv} \equiv \frac{\gamma_e^2 m_e^2 \beta^2}{D_{vv}} \leq$$
$$\leq c_{v,e}^{-1}\beta^3(q-1)\left|1-\frac{1}{2}\left(\frac{v_{\mathrm{th},e}}{v}\right)^2(q+2)\right|^{-1}+$$
$$+c_{v,p}^{-1}\beta^3(q-1)\left|\frac{m_e}{m_p}-\frac{1}{2}\left(\frac{v_{\mathrm{th},p}}{v}\right)^2(q+2)\right|^{-1}. \qquad (31)$$

To explore the dependence on the slope of the suprathermal electron distribution, we estimate the maximum acceleration timescales for the two different regimes as follows:

$$\tau_{vv,\max} \approx \begin{cases} 4\beta^3\left(c_{v,e}^{-1}+\dfrac{m_e}{m_p}c_{v,p}^{-1}\right), \text{for} q = 5, \\ 2\beta^3\left[c_{v,e}^{-1}\left(\dfrac{v}{v_{\mathrm{th},e}}\right)^2+c_{v,p}^{-1}\left(\dfrac{v}{v_{\mathrm{th},p}}\right)^2\right], \\ \text{for} q \to \infty. \end{cases} \qquad (32)$$

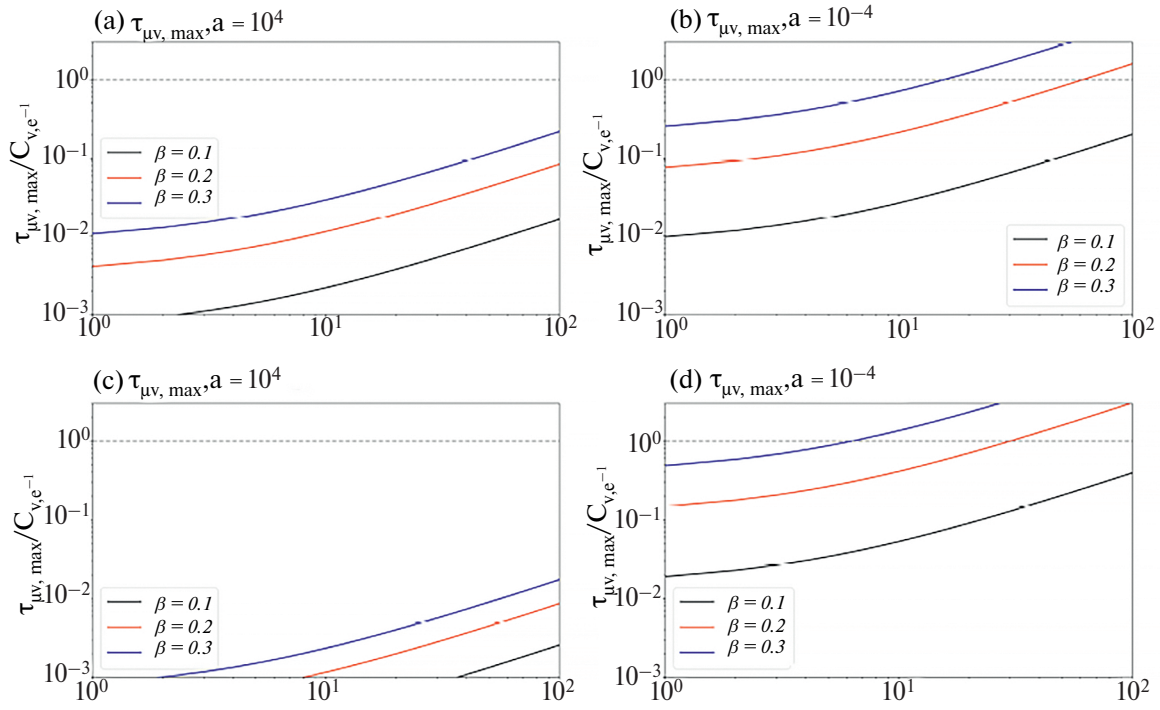

Fig. 3. $\tau_{\mu v,max}$ and $\tau_{\mu\mu,max}$ for weakly (left panels) and strongly (right panels) magnetized plasmas. Here, the electron thermal velocity is assumed as $v_{th,e}/c = 10^{-3}$, and the gray lines display the value $\tau_{max}/c_{v,e}^{-1} = 1$

Because electron velocities satisfy $v/v_{th,e} \gg 1$ and $v/v_{th,p} \gg 1$, the maximum acceleration timescale is much larger when $q \to \infty$. This indicates the evolution of the electron distribution function with a larger q more effectively depends on Coulomb collisions, and such a distribution is likely to resemble a Maxwellian. It is understandable that wave-particle interactions with sufficiently large q are inefficient due to the absence of a sufficient number of resonant particles. Indeed, acceleration timescales become longer regardless of electron velocity for larger q (panel *a* of Fig. 2), and these effects are more pronounced for suprathermal electrons with higher β.

While the analysis in this section has focused on the diagonal terms of the diffusion tensor, it has been demonstrated that the off-diagonal terms, particularly those involving diffusion in pitch angle scattering, are significant in weakly magnetized plasmas. Using equations (19) and (20), we can roughly estimate the maximum values of $\tau_{\mu v}$ and $\tau_{\mu\mu}$ for wave-particle interactions. Applying the inequality (31) to Equations (19) and (20), we obtain

$$\tau_{\mu v} \le \left\{ c_{v,e}^{-1}\beta^3(q-1)\left|1-\frac{1}{2}\left(\frac{v_{th,e}}{v}\right)^2(q+2)\right|^{-1} + \right.$$

$$\left. + c_{v,p}^{-1}\beta^3(q-1)\left|\frac{m_e}{m_p}-\frac{1}{2}\left(\frac{v_{th,p}}{v}\right)^2(q+2)\right|^{-1}\right\} \times$$

$$\times\left|\frac{\beta\mu}{|\mu|}\left(\frac{\beta|\mu|}{a}\right)^{-1/3}+\mu\right|^{-1}, \qquad (33)$$

$$\tau_{\mu\mu} \le \left\{ c_{v,e}^{-1}\beta^3(q-1)\left|1-\frac{1}{2}\left(\frac{v_{th,e}}{v}\right)^2(q+2)\right|^{-1} + \right.$$

$$\left. + c_{v,p}^{-1}\beta^3(q-1)\left|\frac{m_e}{m_p}-\frac{1}{2}\left(\frac{v_{th,p}}{v}\right)^2(q+2)\right|^{-1}\right\} \times$$

$$\times\left|\beta^2\left(\frac{\beta|\mu|}{a}\right)^{-2/3}+\frac{2\beta\mu^2}{|\mu|}\left(\frac{\beta|\mu|}{a}\right)^{-1/3}+\mu^2\right|^{-1}. \qquad (34)$$

Fig. 3 shows the behavior of the two characteristic timescales $\tau_{\mu v}$ and $\tau_{\mu\mu}$ across a wide range of slope parameters q and electron velocities β. In weakly magnetized plasmas $(a = 10^4)$, shown in the left panels of Fig. 3, wave-particle interactions

can effectively transport electrons with softer distribution slopes due to enhanced diffusion in pitch angle space. This enhancement occurs even in scenarios where $\tau_{vv} > c_{v,e}^{-1}$, as $\tau_{\mu\mu} < \tau_{\mu v} \ll c_{v,e}^{-1}$ can be satisfied. Conversely, in strongly magnetized plasmas ($a = 10^{-4}$), shown in the right panels of Fig. 3, diffusion in pitch angle space does not significantly enhance efficient transport through wave-particle interactions when collisional effects dominate ($\tau_{vv} > c_{v,e}^{-1}$), as $\tau_{vv} < \tau_{\mu v} < \tau_{\mu\mu}$.

4. CYCLOTRON RESONANCE OF SUPRATHERMAL ELECTRONS AND NATURE OF WHISTLER WAVES

In this section, we derive the conditions for the minimum velocity of resonant electrons and the characteristics of whistler waves corresponding to wave-particle interaction. The criteria described in this section encompass the characteristics of the turbulent power spectrum, such as its power-law slope, and the effects of Coulomb collisions, as depicted in the schematic Fig. (see Fig. 4). Assuming that the energy transferred through whistler turbulence remains constant across spectra with arbitrary slopes, the maximum wavenumber of a flatter spectrum could be larger than that of a steeper spectrum. Additionally, Coulomb collision effects may suppress energy transport to smaller scales, thereby allowing for a larger maximum wavenumber with stronger Coulomb collisional effects. Such wave characteristics could influence particle transport through turbulence by determining the minimum momentum of electrons required for wave-particle interactions.

Considering only the electron collision term, the minimum velocity criterion can be derived using the inequality (31) as follows:

$$\beta \geq \left[\frac{3}{\pi A}(1-\mu^2)^{-1}|\mu|^{(1-s)/3}\left(\frac{v_{\mathrm{th},e}}{c}\right)^3 \times \right.$$
$$\left. \times (q-1)\left(\frac{\Omega_e^{-1}}{\omega_{pe}^{-1}}\right)^{(2s+1)/3}\left(\frac{c_{v,e}^{-1}}{v_{\mathrm{th},e}^{-3}\omega_{pe}^{-1}}\right)^{-1}\right]^{3/(s+2)}. \quad (35)$$

Here, for simplicity, we consider only electron-electron collisions since the collisional timescales satisfy $c_{v,e}^{-1} \ll c_{v,p}^{-1}$. Clearly, more electrons with lower velocities can be energized through wave-particle interactions when collisional timescales are longer. While the minimum velocity increases as the magnetic field strength decreases (or, $|_e^{-1}$ increases), we interpret that these effects could be minor when considering regimes dominated by wave-particle interactions ($|_e^{-1} \ll c_{v,e}^{-1}$). Additionally, a steeper initial slope of the suprathermal electron distribution q leads to a larger minimum velocity, indicating that transport of suprathermal electrons is less likely when q is sufficiently large.

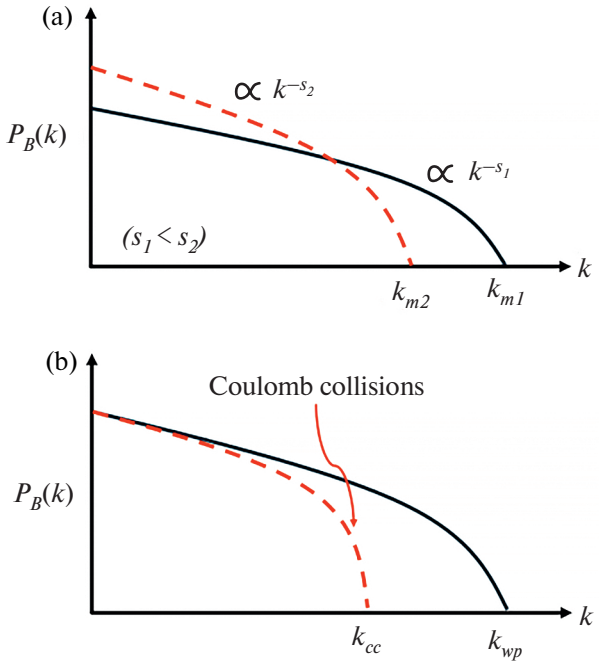

Fig. 4. *a* – Schematic diagrams illustrating whistler turbulence spectra with two different power-law slopes (s_1, s_2). Assuming constant energy transport through whistler turbulence, the maximum wavenumber for a steeper (s_2) spectrum may be smaller than that for a flatter spectrum (s_1) ($k_{m2} < k_{m1}$). *b* – Schematic diagrams demonstrating the influence of Coulomb collisions on turbulent energy transport. Coulomb collisions hinder energy transfer to smaller scales, potentially resulting in a smaller maximum wavenumber (k_{cc}) compared to scenarios without Coulomb collisions (k_{wp})

For low-frequency whistler waves ($\omega_r \ll \Omega_e$), the wavenumber $k_\parallel$ and wavelength $\lambda_\parallel$ for scattering particles are derived as follows:

$$k_\parallel \approx \frac{n\Omega_e}{v\mu}, \quad \lambda_\parallel \equiv \frac{2\pi}{k_\parallel} \approx \frac{2\pi v\mu}{n\Omega_e}. \quad (36)$$

From the inequality (35), we obtain the maximum wavenumber $k_{\parallel,\max}$ and the minimum wavelength $\lambda_{\parallel,\min}$ for wave-particle interactions:

$$\frac{ck_{\parallel,\max}}{\Omega_e} \approx \frac{n}{\mu}\left[\frac{3}{\pi A}(1-\mu^2)^{-1}|\mu|^{(1-s)/3}\left(\frac{v_{\mathrm{th},e}}{c}\right)^3 \times\right.$$

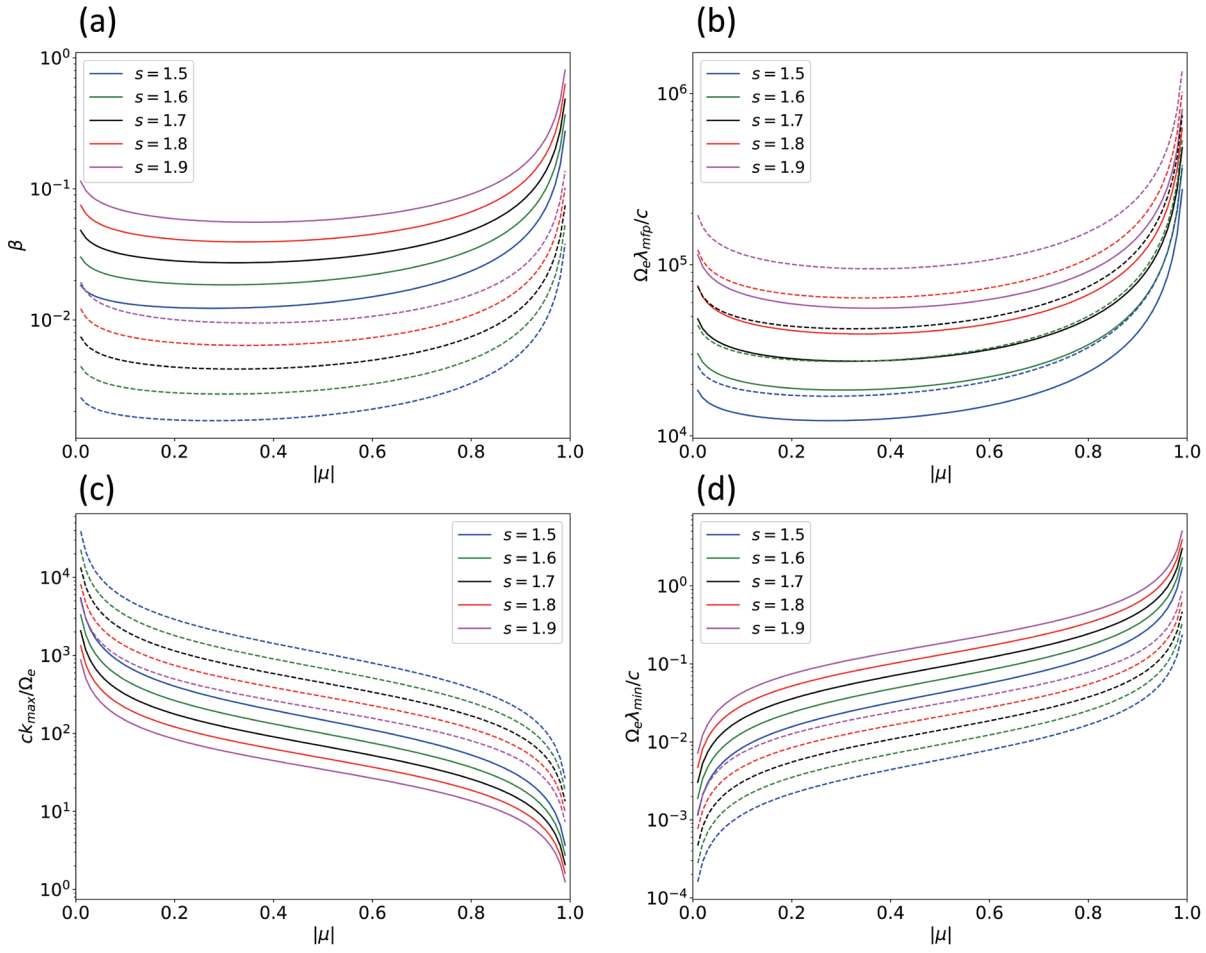

Fig. 5. *a* — Minimum electron velocity, *b* — minimum collisional mean free path, *c* — maximum wavenumber, and *d* — minimum wavelength as functions of pitch angle μ. Solid lines correspond to $c_{v,e}^{-1}/\Omega_e^{-1} = 10^6$, while dashed lines correspond to $c_{v,e}^{-1}/\Omega_e^{-1} = 10^7$. The results are shown for $q = 5$ as an example

$$\times(q-1)\left(\frac{\Omega_e^{-1}}{\omega_{pe}^{-1}}\right)^{(2s+1)/3}\left(\frac{c_{v,e}^{-1}}{v_{\mathrm{th},e}^{-3}\omega_{pe}^{-1}}\right)^{-1}\Bigg]^{-3/(s+2)}, \quad (37)$$

$$\lambda_{\parallel,\min} \approx \frac{2\pi}{k_{\parallel,\max}}. \quad (38)$$

We also consider the minimum collisional length defined as

$$\frac{\Omega_e\lambda_{\mathrm{mfp,min}}}{c} \approx \beta_{\min}\frac{c_{v,e}^{-1}}{\Omega_e^{-1}}. \quad (39)$$

In the criterion (37), the maximum wavenumber decreases as the initial slope of the electron distribution function (q) increases. This indicates that the wavenumber range of wave-particle interactions could be reduced when there are fewer suprathermal electrons (i.e., the spectrum is steeper with larger q).

According to the conditions for resonant scattering and efficient wave-particle interactions, we explore the minimum electron velocity and wave properties relevant to wave-particle interactions across varying power-law slopes of turbulent spectra. The maximum wavenumber $k_{\parallel,\max}$ and minimum wavelength $\lambda_{\parallel,\min}$, derived using the inequality (31) that includes Coulomb collisions and wave-particle interactions, align with the physical insights demonstrated in Fig. 4. Specifically, $k_{\parallel,\max}$ decreases and $\lambda_{\parallel,\min}$ increases as the power-law slope of the turbulent spectra increases. This suggests that turbulence with a flatter spectrum is more efficient at transporting particles. Additionally, as shown by the solid lines in Fig. 5, relatively strong Coulomb collisions can suppress particle transport by reducing $k_{\parallel,\max}$. In contrast, weakly collisional plasmas (represented by dashed lines in Fig. 5) exhibit greater $k_{\parallel,\max}$

values. It is important to note that this analysis generally applies to weakly collisional plasmas where $\lambda_{\parallel,\min} \ll \lambda_{\mathrm{mfp},\min}$.

5. SUMMARY AND DISCUSSION

In this work, we demonstrate how wave-particle interactions through whistler turbulence differ between weakly and strongly magnetized plasmas. In strongly magnetized plasmas (characterized by $a = \omega_{pe}^2 / \Omega_e^2 \ll 1$), the diffusion timescales at large pitch angles ($|\mu| > 0.5$) exhibit saturation for sufficiently small values of a, indicating that strong magnetic fields effectively regulate particle diffusion in pitch angle space. In weakly magnetized plasmas (where $a \gg 1$), on the other hand, large-angle scattering can be enhanced due to the increased magnetization factor a. This enhancement suggests that electron transport via wave-particle interactions may dominate over Coulomb collisions, facilitated by enhanced diffusion in pitch angle space. Additionally, incorporating Coulomb collision effects, we provide conditions for electron transport through whistler turbulence, including the minimum electron velocity and wavelength required for resonant scattering. These findings are broadly applicable to weakly collisional astrophysical plasmas, offering insights into the range of resonant velocities and maximum wavenumbers for wave-particle interactions across a wide range of magnetic field strengths parametrized by a. In such environments, weakly magnetized mediums benefit from efficient transport via wave-particle interactions, particularly when suprathermal particles are present.

We further comment on the significance of investigating particle transport through plasma turbulence in space and astrophysical media. The generation of suprathermal particles is feasible through collisionless shocks or plasma turbulence in various astrophysical environments, with multi-wavelength emissions serving as observational evidence of particle acceleration. While studies on electron transport via whistler turbulence have predominantly focused on non-Maxwellian electron distributions in solar wind environments, similar investigations in diverse astrophysical contexts are warranted. For example, research has shown that velocity anisotropy in interstellar and intracluster media can induce whistler waves [27, 40, 41], potentially maintaining non-Maxwellian electron distributions within localized regions experiencing whistler turbulence. Additionally, it has been shown that suprathermal electrons can be generated by various plasma instabilities in astrophysical media, including whistler, firehose, mirror, and cyclotron instabilities. In particular, current drive exhibited in localized areas, such as the upstream and downstream regions of collisionless shocks, could trigger plasma instabilities that significantly amplify the magnetic field and generate suprathermal particles through waves satisfying cyclotron resonance conditions [26–28, 40–43]. The characteristics of these plasma instabilities and their acceleration efficiency depend on the properties of collisionless shocks, including the shock Mach number, plasma magnetization, and the geometry of the background magnetic field [27, 40]. Moreover, Lower-Hybrid waves could be induced by diamagnetic currents in inhomogeneous plasma systems, which typically propagate in space and astrophysical plasmas, including those with compressible turbulence. The roles of particle acceleration or heating through Lower-Hybrid waves have also been proposed [44, 45]. In this context, it is necessary to conduct further investigations, including the theory of particle transport through various plasma instabilities triggered in astrophysical media, corresponding numerical simulations to support the theory, and complementary observations representing particle acceleration and heating.

REFERENCES

1. *R. Dumont*, Encyclopedia of Nuclear Energy **3**, 479 (2021).
2. *A. A. Schekochihin, S. C. Cowley, and W. Dorland*, Plasma Physics and Controlled Fusion **49**, A195 (2007).
3. *C. Vogt and T. A. Ensslin*, Astronomy and Astrophysics **434**, 67 (2005).
4. *A. H. Minter and S. R. Spangler*, Astrophys. J. **458**, 194 (1996).
5. *R. J. Leamon, C. W. Smith, N. F. Ness, W. H. Matthaeus, and H. K. Wong*, J. Geophys. Res. **103**(A3), 4775 (1998).
6. *S. Regnier*, Astronomy and Astrophysics **581**, A9 (2015).
7. *S. Ni, Y. Chen, C. Li, Z. Zhang, H. Ning, X. Kong, B. Wang, and M. Hosseinpour*, Astrophys. J. Lett. **891**, L25 (2020).
8. *B. Cerutti and G. Giacinti*, Astronomy and Astrophysics **642**, A123 (2020).
9. *C. F. Kennel and F. V. Coroniti*, Astrophys. J. **283**, 694 (1984).

10. *A. A. Schekochihin, S. C. Cowley, W. Dorland, G. W. Hammett, G. G. Howes, E. Quataert, and T. Tatsuno*, Astrophys. J. Supplement Series **182**, 310 (2009).

11. *G. G. Howes*, ArXiv:2402.12829 (2024).

12. *L. Bercic, M. Maksimovic, J. S. Halekas et al.*, Astrophys. J. **921**, 83 (2021).

13. *J. Halekas, P. Whittlesey, D. Larson et al.*, Astrophys. J. Supplement Series 246, 22 (2020).

14. *S. Kim, P. H. Yoon, G. Choe, and L. Wang*, Astrophys. J. **806**, 32 (2015).

15. *P. H. Yoon*, J. Phys.: Conference Series **642**, 012030 (2015).

16. *B. Tang, L. Adhikari, G. P. Zank, and H. Che*, Astrophys. J. **964**, 180 (2024).

17. *B. Tang, G. P. Zank, and V. I. Kolobov*, Astrophys. J. **924**, 113 (2022).

18. *V. Pierrard, M. Lazar, and R. Schlickeiser*, Solar Physics **269**, 421 (2011).

19. *G. P. Zank*, *Transport Processes in Space Physics and Astrophysics*, Springer, Berlin (2014).

20. *O. Allanson, T. Elsden, C. Watt, and T. Neukirch*, Frontiers in Astronomy and Space Sciences **8**, 805699 (2022).

21. *S.-Y. Jeong, D. Verscharen, R. T. Wicks, and A. N. Fazakerley*, Astrophys. J. **902**, 128 (2020).

22. *J. Steinacker and J. A. Miller*, Astrophys. J. **393**, 764 (1992).

23. *C. Vocks, C. Salem, R. P. Lin, and G. Mann*, Astrophys. J. **627**, 540 (2005).

24. *A. Micera, Zhukov A. N., R. Lopez et al.*, Astrophys. J. Lett. **903**, L23 (2020).

25. *A. Micera, Zhukov A. N., R. Lopez et al.*, Astrophys. J. **919**, 42 (2021).

26. *D. Caprioli and A. Spitkovsky*, Astrophys. J. **783**, 91 (2014).

27. *J.-H. Ha, S. Kim, D. Ryu, and H. Kang*, Astrophys. J. **915**, 18 (2021).

28. *A. Bohdan, J. Niemiec, O. Kobzar, and M. Pohl*, Astrophys. J. **847**, 71 (2017).

29. *V. Bresci, M. Lemoine, L. Gremillet, L. Comisso, L. Sironi, and C. Demidem*, Phys. Rev. D1**06**, 023028 (2022).

30. *L. Comisso and L. Sironi*, Astrophys. J. Lett. **936**, L27 (2022).

31. *C. Vega, S. Boldyrev, V. Roytershteyn, and M. Medvedev*, Astrophys. J. Lett. **924**, L19 (2022).

32. *P. Helander and D. J. Sigmar*, *Collisional Transport in Magnetized Plasmas*, Cambridge Univ. Press, Cambridge (2005), Vol. 4.

33. *P. H. Yoon*, Phys. Plasmas **19**, 052301 (2012).

34. *P. H. Yoon, J. Geophys*. Res.: Space Physics **119**, 7074 (2014).

35. *V. Pierrard, M. Lazar, and S. Stverak*, Frontiers in Astronomy and Space Sciences **9**, 892236 (2022).

36. *N. J. Fisch*, Rev. Mod. Phys. **59**, 175 (1987).

37. *S. I. Popel and V. N. Tsytovich*, Contributions to Plasma Physics **31**, 77 (1991).

38. *O. Chang, S. P. Gary, and J. Wang*, Phys. Plasmas **21**, 052305 (2014).

39. *E. Camporeale and G. Zimbardo*, Phys. Plasmas **22**, 092104 (2015).

40. *S. Kim, J.-H. Ha, D. Ryu, and H. Kang*, Astrophys. J. **913**, 35 (2021).

41. *M. A. Riquelme and A. Spitkovsky*, Astrophys. J. **733**, 63 (2012).

42. *A. R. Bell*, Monthly Notices of the Royal Astronomical Society **353**, 550 (2004).

43. *E. Amato and P. Blasi*, Monthly Notices of the Royal Astronomical Society **392**, 1591 (2009).

44. *I. Shinohara and M. Hoshino*, Adv. Space Res. **24**, 43 (1999).

45. *F. Lavorenti, P. Henri, F. Califano, S. Aizawa, and N. Andre*, Astronomy and Astrophysics **652**, A20 (2021).

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS, 2025, Vol. 167, No. 1, pp. 136–142

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

RESONANT WAVEGUIDE LUMINESCENCE LOSSES IN A LIQUID CRYSTAL LAYER CONFINED BY ITO ELECTRODES

© 2025 S. P. Palto, D. O. Rybakov, A. R. Geivandov*, I. V. Kasyanova

Shubnikov Institute of Crystallography, Kurchatovsky Complex "Crystallography and Photonics" National Research Centre Kurchatov Institute, Moscow, Russia

**e-mail: ageivandov@yandex.ru*

Received August 15, 2024

Revised September 05, 2024

Accepted September 05, 2024

Abstract. In a planarly-aligned layer of a nematic liquid crystal (LC) with a luminescent dye, the spectra of luminescence arising upon laser excitation of dye molecules and propagating in the waveguide mode were investigated. It was shown that the presence of ITO electrodes confining the LC layer leads to significant resonant losses of radiation energy. These losses are explained by phase synchronism between the waveguide modes in the LC layer and the ITO electrodes. The spectral position of the loss maxima depends on the polarization state of light, and their intensity increases with decreasing LC layer thickness. It was shown that the use of LC alignment layers made of fluorinated polymers with a low refractive index coated onto ITO electrodes allows one to significantly suppress resonant radiation losses.

Keywords: *liquid crystal, ITO losses, waveguide mode, luminescence, LC lasing*

DOI: 10.31857/S00444510250112e6

1. INTRODUCTION

Alongside the widespread use of liquid crystals (LCs) in information display technologies, there has been increasing interest in utilizing LCs for various photonic devices. In particular, LC materials doped with laser dyes can be employed both for studying the photonic properties of LCs [1, 2] and as an active medium for microlaser systems [3, 4, 5, 6, 7]. Among these systems, microlasers operating in the waveguide light generation mode hold a special place [8, 9, 10, 11, 12]. However, despite the many advantages of the waveguide generation mode, this approach also has certain drawbacks. For example, controlling the LC layer to create spatially periodic refractive index modulation and, consequently, distributed feedback, requires control electrodes that confine the LC layer. In LC devices, transparent electrodes based on indium tin oxide (ITO) are widely used. In this case, an inevitable problem arises due to the need to minimize light energy losses in the electrodes during waveguide mode propagation.

In [13], numerical FDTD modeling demonstrated that light propagation in the waveguide mode within oriented LC layers confined by transparent ITO electrodes is characterized by significant resonant losses. These losses occur in specific spectral ranges due to phase-synchronous energy transfer from the liquid crystal layer to the thin electrode layers. The present study aims to experimentally observe the resonant losses predicted in [13] by exciting luminescence in the LC layer and recording the luminescence spectrum at the output of the liquid crystal waveguide formed by the LC layer and the confining layers, including ITO. The study also investigates the possibility of reducing these losses by introducing alignment layers with a low refractive index between the ITO electrodes and the LC layer, as recommended in [13].

2. EXPERIMENTAL SAMPLES

The experimental scheme of the liquid crystal (LC) cell with ITO electrodes is shown in Fig. 1. The cell consists of two glass substrates 1, 2, with transparent ITO electrodes 3, 4 at their inner sides.

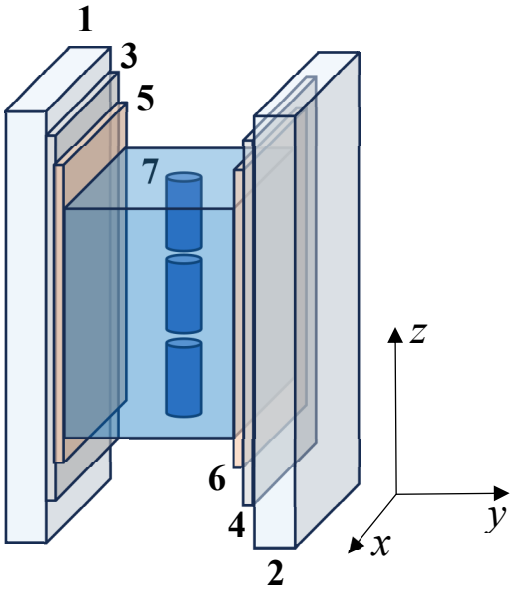

Fig. 1. Schematic diagram of the layered structure of the LC cell. *1*, *2* – glass substrates; *3*, *4* – ITO electrode layers; *5*, *6* – polymer alignment layers rubbed in the direction *z*; *7* – LC layer (E7) with DCM dye (cylinder axes indicate the LC director direction).

We used industrial glass for display technologies with a measured ITO electrode thickness of 150 ± 10 nm. To achieve planar alignment of the nematic LC (E7, Merck), thin polymer films 5, 6 were applied to the ITO surfaces and mechanically rubbed with a soft cloth along the *z*-axis, determining the easy axis direction and, consequently, the optical axis direction in the LC layer. Two types of polymers were used for the alignment films: (a) polyimide (PI) with a refractive index of 1.65 (AD9103 lacquer, NPO Plastik) and (b) fluorinated polymer with a refractive index of 1.42 (copolymer of tetrafluoroethylene and vinylidene fluoride, F42-V). The LC layer thickness, alignment film type, and presence of ITO electrodes varied depending on the sample number (see table).

As seen from the table, Sample 1 does not contain ITO electrodes. This LC cell was used as a reference sample to visualize spectral changes in the emitted light due to the presence of ITO electrodes.

The choice of the liquid crystal E7 is due to the extensive study of this material and the availability of many of its physical parameters. For example, the spectral dependencies of the refractive index, crucial for our studies, are well known across a wide

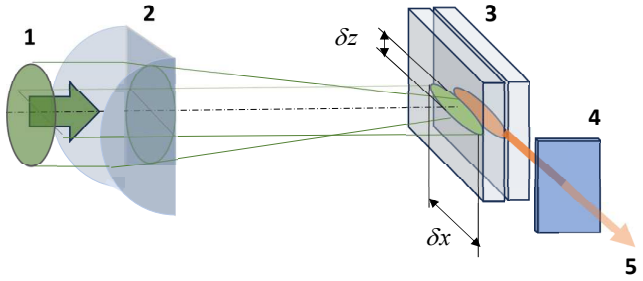

Fig. 2. Scheme of luminescence excitation in the LC layer. *1* – laser beam (λ = 532 nm, τ = 10 ns); *2* – cylindrical lens (focal length *f* = 100 mm, characteristic focusing area dimensions δ*z* = 0.1 mm; δ*x* = 3 mm); *3* – liquid crystal cell; *4* – set of optical filters and polarizer; *5* – radiation registered by a fiber optic spectrometer.

spectral range [14], allowing us to build a realistic model of resonant losses in the E7 layer between ITO electrodes in [13]. To impart luminescent properties to the LC layer, we used the well-known laser dye DCM (4-(Dicyanomethylene)-2-methyl-6-(4-dimethylaminostyryl)-4H-pyran, Sigma Aldrich, 0.6 wt.%). This dye is characterized by intense luminescence in the 570–650 nm wavelength range and is widely used to achieve lasing effects in various LC systems.

The luminescence excitation and registration scheme is shown in Fig. 2. Luminescence excitation in the LC layer was performed using radiation 1 from a neodymium laser operating in Q-switched mode at a wavelength of λ=532 nm with a pulse duration of 10 ns. The pulse energy was approximately 80 μJ. The laser radiation was linearly polarized along the direction z (along the LC director), ensuring maximum luminescence efficiency [11]. The laser beam was focused on the LC layer in cell 3 by cylindrical lens 2 into a narrow stripe with a width of $\delta z = 0.1$ mm and a length of $\delta x = 3$ mm along the direction x of waveguide luminescence propagation 5. The position x_0 of the focused area center, measured from the LC cell end, varied from 1.5 to 2.5 mm.

Luminescence from the LC layer end face was recorded using a fiber optic spectrometer Avantes Avaspec 2048. To exclude the registration of light propagating into the substrates, the ends of the latter were coated with an opaque (black) dye layer, and a mask with a slit was used. To register polarization spectra, a polarizer 4 was placed in

Table. Analyzed samples and their parameters

Sample No.	LC layer thickness, μm	Sample type	Alignment layer thickness, nm	ITO presence
1	6.7 ± 0.2	PI	20 ± 10	No
2	6.8 ± 0.2	PI	20 ± 10	Yes
3	2.4 ± 0.2	PI	20 ± 10	Yes
4	12 ± 0.2	PI	20 ± 10	Yes
5	6.3 ± 0.2	Ф42-В	350 ± 10	Yes

front of the fiber optic cable lens, allowing the recording of TE-polarized spectra (electric field oscillations along the z axis and LC director, see Fig. 1) and TM-polarized spectra (electric field oscillations in the xy plane). Additionally, glass optical filters were used to attenuate both scattered laser radiation and luminescence when necessary, installed alongside the polarizer in front of the fiber optic cable input lens of the spectrometer.

3. DISCUSSION OF RESULTS

Fig. 3 shows the spectra of unpolarized luminescence for the reference sample 1 (see table) without ITO electrodes (curve 1) and sample 2 (curve 2), obtained under identical laser excitation pulse energies (approximately 80 μJ). It is worth noting that the luminescence intensity here and below is presented on a logarithmic scale. As seen, the luminescence intensity for sample 2 with ITO electrodes is significantly lower than that recorded for the reference sample 1. In the spectrum (curve 2), a characteristic dip at the wavelength of 588 nm is observed, which is absent in the sample without ITO. When the spectrum of sample 2 is divided by the spectrum of sample 1, the spectral dependence of the relative losses I_2/I_1 in sample 2 compared to sample 1 is obtained (see the inset in Fig. 3).

It is evident that the relative losses associated with the presence of ITO electrodes are characterized by a spectral band with a maximum absorption at the wavelength of 592 nm. The luminescence intensity at this wavelength for sample 2 is approximately 8 times lower than that for sample 1. There is also an increase in losses at wavelengths longer than 625 nm. Unfortunately, it is challenging to register this longer-wavelength band accurately across the entire range due to the very low luminescence intensity at wavelengths above 650 nm.

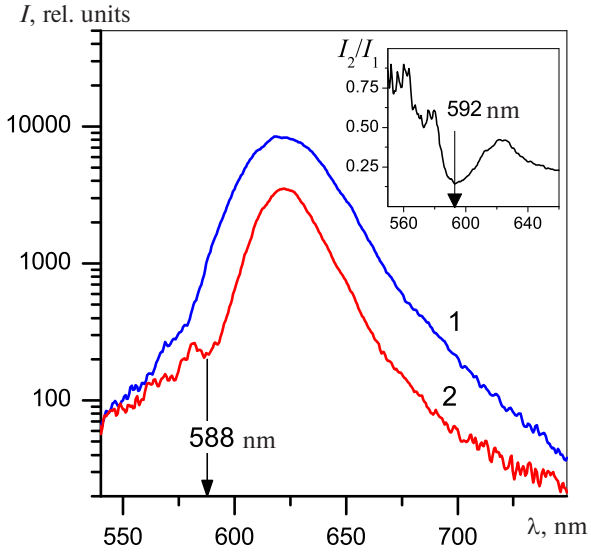

Fig. 3. Luminescence spectra at the output of the LC cell (see Fig. 2) after light propagation in the LC layer in the waveguide mode. The pump area length determining the propagation distance is $\delta x = 3$ mm, and the distance from the pump center to the LC cell edge is $x_0 = 1.5$ mm. Curve *1* – spectrum of sample 1 (no ITO electrodes); curve *2* – spectrum of sample 2 (with ITO electrodes, thickness 150 nm). The inset shows the ratio of the spectrum of sample 2 to that of sample 1.

According to the numerical calculations in [13] for a planarly-aligned liquid crystal layer E7 confined by ITO electrodes with a thickness of 170 nm, two resonance bands exist in the spectral range of 550–900 nm, with maximum losses at wavelengths of волн $\lambda_1 = 570$ nm and $\lambda_2 = 705$ nm for TE- and TM-polarized light, respectively. The spectral position of these losses does not depend on the thickness of the LC layer. However, as shown in [13], changes in the ITO layer thickness and

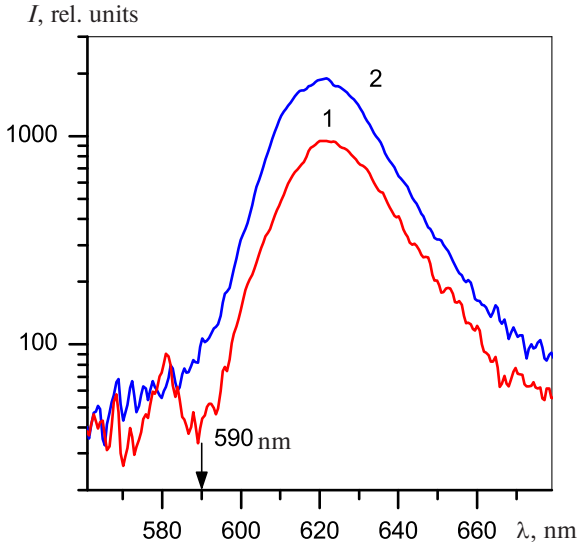

Fig. 4. Polarization spectra of luminescence measured for sample 1, where curve *1* corresponds to TE-polarization and curve *2* to TM-polarization. The distance from the pump center to the LC cell edge is $x_0 = 1.5$ mm.

the presence of a polyimide alignment film can shift the spectral position of the resonance bands. Considering the experimental error associated with measuring the ITO layer thickness and the presence of a thin polyimide alignment film in experimental sample 2, we associate the observed loss maximum at 592 nm with the resonance band calculated at $\lambda_1 = 570$ nm for the TE-polarized mode in [13]. Similarly, the increasing losses at wavelengths above 625 nm (inset in Fig. 3) are explained by the calculated resonance band at $\lambda_2 = 705$ nm for TM-polarized light. The spectral data in Fig. 4, showing the polarization spectra of luminescence, confirm this conclusion. The luminescence intensity dip at 588 nm is characteristic only for the TE-polarized mode (curve 1 in Fig. 4). As the wavelength increases beyond 625 nm, the intensity of TM-polarized luminescence, shown by curve 2, decreases faster than that of the TE mode (curve 1). Thus, the observed long-wavelength losses also agree with the numerical model in [13].

According to the analytical model of a thin ITO layer with a refractive index n_1 between the glass substrate n_0 and the liquid crystal layer n_2, the wavelengths corresponding to the maxima of the resonance losses are determined by the following relations [13]:

$$\lambda_{m,TE,TM} = \frac{2d_{ITO}n_2}{m - \frac{\delta\phi_{TE,TM}}{2\pi}} \times \sqrt{\left(\frac{n_1}{n_{2,TE,TM}}\right)^2 - 1}, \quad (1)$$

where the indices TE and TM refer to TE- and TM-polarized light, respectively, and m is a natural number. The additional phase shifts $\delta\phi$ in Equation (1), associated with double reflection of waves in the ITO layer from the boundaries of the ITO-glass substrate and the ITO-LC layer interfaces, are determined for TE- and TM-polarized light by the following expressions:

$$\delta\phi_{TE} = -2\left[arctg\left(\frac{\sqrt{\sin^2(\theta_{1,TE}) - (\frac{n_0}{n_1})^2}}{\cos(\theta_{1,TE})}\right) + arctg\left(\frac{\sqrt{\sin^2(\theta_{1,TE}) - (\frac{n_{2,TE}}{n_1})^2}}{\cos(\theta_{1,TE})}\right)\right],$$

$$\delta\phi_{TM} = 2\pi - 2\left[arctg\left(\frac{\sqrt{\sin^2(\theta_{1,TM}) - (\frac{n_0}{n_1})^2}}{(\frac{n_0}{n_1})^2\cos(\theta_{1,TM})}\right) + arctg\left(\frac{\sqrt{\sin^2(\theta_{1,TM}) - (\frac{n_{2,TM}}{n_1})^2}}{(\frac{n_{2,TM}}{n_1})^2\cos(\theta_{1,TM})}\right)\right].$$

It is also important to consider the spectral dispersion of refractive indices: $n_0 \equiv n_0(\lambda)$ is the refractive index of the glass substrate, $n_1 \equiv n_1(\lambda)$ is the refractive index of ITO, $n_{2,TE,TM} \equiv n_{2,TE,TM}(\lambda)$ is the polarization-dependent refractive index of the liquid crystal (for planarly-aligned LC, $n_{2,TE} = n_{\parallel}$, $n_{2,TM} = n_{\perp}$), θ_1 is the angle between the layer normal and the wave vector in the ITO layer (for phase-synchronized coupling of the planar mode from the LC layer to the ITO, $\sin\theta_1 = n_1 / n_{2,TE,TM}$).

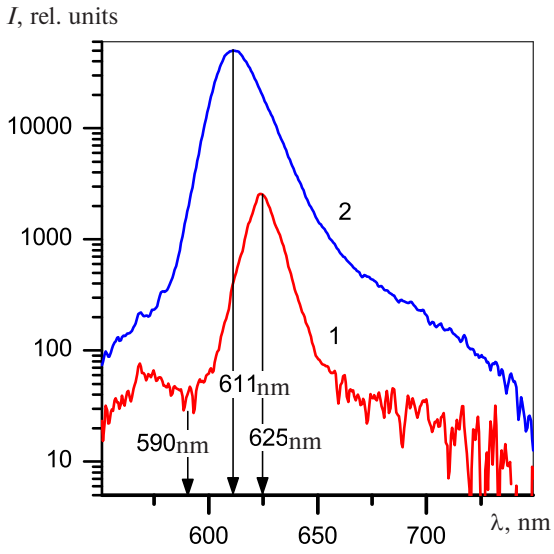

Fig. 5. Spectra of unpolarized luminescence for sample 3 (curve *1* $d = 2.4$ μm) and sample 4 (curve *2*, $d = 12$ μm). Distance from the pump center to the LC cell edge is $x_0 = 2.5$ mm.

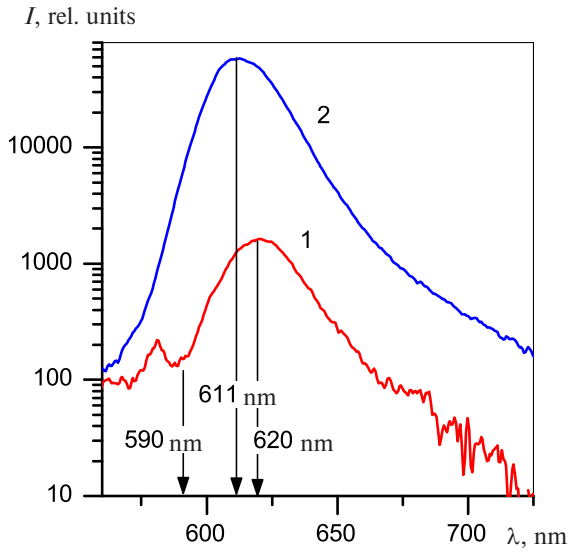

Fig. 6. Spectra of unpolarized luminescence for sample 2 (curve *1*, $d = 6.8$ μm) and sample 5 (curve *2*, $d = 6.3$ μm). Distance from the pump center to the LC cell edge $x_0 = 2.5$ mm.

As follows from Equation (1), real solutions exist only under the condition $n_1 \geq n_2$. For ITO, there is strong spectral dispersion of the refractive index [13], so the corresponding solutions exist only in specific spectral ranges, which differ for TE- and TM-polarized light due to the optical anisotropy of the liquid crystal and, accordingly, the conditions:

$$n_1 \geq n_{2,TE}, \quad n_1 \geq n_{2,TM}.$$

As shown in [13], for planar LC alignment, there are only two solutions for TM-polarized light: $\lambda_m \cong 720$ nm for $m = 1$ and $\lambda_m \cong 440$ nm for $m = 2$. Since the luminescence spectrum is limited to approximately 550–700 nm, we can observe only the short-wavelength edge of the TM mode absorption for $m = 1$, which appears at wavelengths above 625 nm (see the inset in Fig. 3).

For TE polarization, the corresponding loss peak occurs at $\lambda = 570$ nm, which, considering measurement errors and differences between our experiment and the model, is very close to the observed peak at 592 nm (see the inset in Fig. 3). In the experiment, as in the model, this loss peak is observed exclusively for TE-polarized light (Fig. 4).

The model in [13] predicts a significant increase in losses with decreasing LC layer thickness. This was confirmed experimentally (Fig. 5). Here, curve 1 corresponds to sample 3 (see table) with an LC layer thickness of $d = 2.4$ μm, and curve 2 corresponds to a thickness of $d = 12$ μm. Both curves are for unpolarized light, showing the loss peak at 592 nm and the onset of a sharp luminescence decrease at wavelengths above 625 nm, associated with the existence of a longer-wavelength loss band peaking beyond 700 nm. A comparison of luminescence intensities at 590 nm reveals that reducing the thickness from 12 to 2.4 μm increased the losses by approximately 40 times. The presence of strong loss bands for sample 3 around $\lambda_m \cong 590$ nm and in the longer-wavelength region ($\lambda_m > 700$ nm) leads to significant narrowing of the luminescence spectrum (curve 1, Fig. 5), with the luminescence maximum shifting to the longer-wavelength region near $\lambda = 625$ nm, where losses are minimal. In sample 4 (curve 2), the loss bands appear only as shoulders in the luminescence spectrum, and there is virtually no shift in the luminescence peak ($\lambda = 611$ nm). Notably, at a fixed pump energy of approximately 80 μJ, the luminescence intensity peak in sample 4 was so high that we had to shift the pump center from the LC cell edge to $x_0 = 2.5$ mm to remain within the dynamic range of the spectrometer.

Thus, the spectral measurements fully confirmed the presence of resonance losses caused by the ITO electrodes. According to the aforementioned numerical modeling, resonance losses can be significantly suppressed by introducing thin

low-refractive-index films between the LC layer and the ITO electrodes. This condition is satisfied in sample 5, where relatively thick (350 nm) fluorinated polymer F42-V films with a refractive index of 1.42 were used as LC alignment layers. The results of luminescence spectra measurements were quite impressive (Fig. 6). At a fixed pump pulse energy of 80 μJ, the peak luminescence intensity increased by approximately 50 times compared to sample 2. The characteristic loss band with a maximum at 590 nm disappeared, as did the long-wavelength spectral losses characteristic of sample 2. At wavelengths above 650 nm, luminescence increased significantly, so that even at 750 nm, the measured luminescence signal was significantly above the noise level.

To note, the luminescence intensity in sample 5 significantly exceeds not only that of sample 2 but also that of sample 1, where ITO electrodes are absent. Thus, the high luminescence intensity in sample 5 is not only due to the elimination of resonance losses. We hypothesize that another significant factor contributing to the increased luminescence intensity in sample 5 is the low refractive index of the fluoropolymer ($n = 1.42$), which is significantly lower than that of the glass substrates ($n_0 = 1.51$). As a result, a significantly larger number of TE- and TM-polarized waveguide modes can propagate in the LC layer of sample 5 compared to samples 1 and 2. Indeed, in samples 1 and 2, the lowest refractive index of the LC, $n_\perp = 1.52$, interacting with TM-polarized modes, is very close to the refractive index of the display glass (1.51). Therefore, the critical angle relative to the substrate plane, below which waveguide modes exist, is very small, causing a significant amount of TM-polarized luminescence propagating at angles above the critical angle to leak into the glass substrates.

The situation is further complicated by the fact that, in reality, the LC director does not strictly coincide with the TE polarization direction due to a slight (2–4°) pretilt angle of the director relative to the substrate plane. Thus, even TE-polarized radiation, for which the waveguide condition is satisfied over a wide range of propagation angles, partially converts into TM-polarized modes that leak into the substrate. Resonance losses, in turn, are characterized by relatively broad spectra, significantly reducing luminescence intensity even at wavelengths far from the resonance maxima. This is evident not only from the inset in Fig. 3 but also, for example, in Fig. 5 for sample 3 (curve 1), where, as already noted, significant losses on the "tails" of the resonance bands lead to spectral narrowing and a shift of the luminescence maximum. Thus, resonance losses lead to reduced luminescence across the entire spectral range. It is also worth noting that the modeling in [13], where the resonance bands are relatively narrow, was performed for a single-mode regime, where the light "injected" into the waveguide was characterized by a wave vector strictly parallel to the LC layer plane.

4. CONCLUSION

The experiment confirmed the presence of resonance losses during light propagation in waveguide mode within an LC layer confined by ITO electrodes. The observed spectral loss bands are polarization-sensitive. The spectral position of these bands does not depend on the LC layer thickness, while their intensity increases with decreasing layer thickness.

It was also demonstrated that the use of fluorinated polymer alignment layers with a low refractive index effectively suppresses resonance losses. The obtained results are significant for the application of waveguide modes in electrically controlled LC devices utilizing light propagation in waveguide mode, particularly for liquid crystal microlasers.

FUNDING

This work was carried out as part of the state assignment of the National Research Center "Kurchatov Institute."

REFERENCES

1. *P. V. Dolganov, V. K. Dolganov*. JETP Lett. **108**, 170 (2018).
2. *P. V. Dolganov*. JETP Lett. **105**, 657 (2017).
3. *I. P. Il'chishin, E. A. Tikhonov, V. G. Tishchenko et al.* JETP Lett. **32**, 24 (1980).
4. *W. Cao, A. Munos, P. Palffy-Muhoray et al*. Nature Mater. **1**, 111 (2002).
5. *A. Chanishvili, G. Chilaya, G. Petriashvili et al*. Appl. Phys. Lett. **86**, 051107 (2005).

6. *J. Ortega, C. L. Folcia, J. Etxebarria*. Materials **11**, 5 (2018).

7. *S. P. Palto, N. M. Shtykov, B. A. Umanskii et al*. J. Appl. Phys. **112**, 013105 (2012).

8. *T. Matsui, M. Ozaki, K. Yoshino*. in Proc. SPIE5518, Liquid Crystals VIII (15 October 2004).

9. *Y. Inoue, H. Yoshida, K. Inoue et al*. Appl. Phys. Express **3**, 102702 (2010).

10. *H. Yunxi, Z. Xiaojuan, Y. Benli et al*. Nanophotonics **10**, 3541 (2021).

11. *N. M. Shtykov, S. P. Palto, B. A. Umanskii et al*. Crystallography Rep. **64**, 305 (2019).

12. *N. M. Shtykov, S. P. Palto, A. R. Geivandov et al*. Opt. Lett. **45**, 4328 (2020).

13. *S. P. Palto, A. R. Geivandov*. Photonics **10**, 1089 (2023).

14. *J. Li, C.-H. Wen, S. Gauza et al*. J. of Display Technol. **1**, 52 (2005).