Zn Atoms in Lithium Niobate and Mechanism of Their Insertion into Crystals

T. S. Chernaya, B. A. Maksimov, T. R. Volk, N. M. Rubinina*, and V. I. Simonov

Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninskiĭ pr. 59, Moscow, 117333 Russia
* Moscow State University, Vorob'evy gory, Moscow, 119899 Russia
Received November 21, 2000; in final form, December 21, 2000

Precision X-ray structural studies were carried out for $LiNbO_3 : Zn_x$ single crystals with $x = 0.0$, 2.87, 5.20, and 7.60 at. %. It was found that the insertion of the Zn atoms into the Li position was accompanied by a decrease in the concentration of intrinsic Nb_{Li} defects. At $x = 7.6\%$, the Zn atoms change their locations in the lattice and partially occupy the Nb positions. This clarifies the structural nature of the "threshold" Zn concentration, which manifests itself as singularities in the concentration dependences of various optical properties. The structural origin of the threshold concentration is likely a common feature of all nonphotorefractive impurities (Mg, Zn, In, and Sc) in $LiNbO_3$. A change in the intrinsic defect structure of the $LiNbO_3$ crystals with different Zn concentrations is discussed. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 61.72.Dd; 61.72.Ss; 61.66.Fn; 61.10.Eq

Lithium niobate is a universal material for optical applications. Of considerable interest are $LiNbO_3$ crystals containing the Mg, Zn, In, and Sc impurities named "nonphotorefractive" or "optical-damage resistant" (see bibliography in reviews [1, 2]) because they suppress photorefraction by no less than two orders of magnitude at certain doping levels and, thereby, enhance crystal tolerance to intense laser radiation. Doping with these ions also modifies some optical properties of $LiNbO_3$. Numerous studies have led to the conclusion that the main role of the ions of these type amounts to controlling the concentration of intrinsic (structural) defects in $LiNbO_3$ [1, 2]. Recall that intrinsic defects arise in $LiNbO_3$ crystals as a result of the deviation from the stoichiometry caused by the Li deficiency [1]. According to the presently adopted model that was substantiated by the structural studies [3–5], some of the Li vacancies are occupied by Nb atoms, which typically form the $Nb_{Li}^{4\bullet}$ structural defects.[1] Their charges are compensated by either V'_{Li} [4, 6] or $V_{Nb}^{5'}$ [1, 3]. This model implies that the nonphotorefractive ions, when they occupy the Li position, reduce the concentration of the intrinsic Nb_{Li} defects and, correspondingly, the concentration of their compensators. The location of Mg in the Li position is proved by structural measurements [7]. A decrease in the Nb_{Li} concentration is precisely the factor which was assumed to be the main reason why the effect on the optical properties of $LiNbO_3$ is common to all nonphotorefractive impu-

rities [1, 2]. The commonness of the effect of these impurities is clearly demonstrated by the occurrence of singularities (extremes, kinks, or breaks) in the concentration curves for different optical properties (e.g., maxima in the concentration dependences of 90° phase matching in SHG, parallel shifts of IR bands, etc.) at certain "threshold" concentrations. The threshold concentrations are close for the bivalent Mg and Zn ions (5–6 and 7–8 at. %, respectively) and tervalent In and Sc ions (1.5–2%). There is some debate over the nature of the threshold concentrations [1, 2, 8–10]; in particular, it was speculated that they are associated with a change in the location of dopants in the lattice [1, 2, 8]. However, so far all structural models for the insertion of nonphotorefractive ions into the $LiNbO_3$ lattice and the discussion about the relation between the threshold concentrations and changes in the defect structure have been based only on chemical analysis data [9, 10] and spectroscopic tests [1, 2] without structural studies.

Our precision X-ray structural study was aimed at determining the location of the Zn atoms for different concentrations in the $LiNbO_3$ lattice. The main purpose of this work was to trace the influence of Zn on the intrinsic defect structure of $LiNbO_3$ and analyze the structural origin of the threshold concentration. The choice of impurity was dictated by the fact that the highest threshold Zn concentration (~7%) provided the most reliable estimates.

Single-domain crystals were grown by the Czochralski method from a congruent melt ([Li_2O]/[Nb_2O_5] = 48.5 : 51.5 mol %); zinc was added to the melt in the form of oxide. Samples with zinc contents of 0.0, 2.87, 5.20, and 7.60 at. % in the crystal

---

[1] The Kröger–Winck notations are used: a heavy point and a prime label the defects charged, respectively, positively and negatively with respect to the lattice.

were studied. The method of preparing single-domain samples is described in [11].

X-ray experiments were carried out for samples shaped like spheres with diameters ranging from 0.15(1) to 0.28(1) mm. Measurements were accomplished on an Enraf-Nonius CAD-4F diffractometer using Mo$K\alpha$ radiation. Diffraction data were collected over a full reflection sphere of the reciprocal space for the scattering angles constrained to $2\vartheta \leq 100°$. Intensity arrays containing from 3559 to 5092 reflections with $I \geq 3\sigma(I)$ were obtained for different samples. Averaging of the symmetrically equivalent reflections gave working arrays containing from 640 to 943 unique structure amplitudes. After averaging, the discrepancy factors did not exceed 2.2%. The maximum number of refined structural parameters did not exceed 28 even if the anisotropy and anharmonicity of atomic thermal vibrations were taken into account. All calculations were performed using the Prometheus program package [12].

The structure of $LiNbO_3$ is realized in the $R3c$ space group with unit-cell parameters $a = 5.149(1)$ and $c = 13.862(2)$ Å of the matrix crystal in hexagonal basis. The structural models were refined by the full-matrix least-squares method with a step-by-step scan if there was a strong correlation between the refined parameters [13]. The correlations considerably hampered the reliable determination of site occupancies for atoms. The final discrepancy factors between the calculated and experimental moduli of structure amplitudes were 0.90–1.62% for the unweighed and 0.79–1.38% for the weighed R factors. Selected structural crystallographic

data for the single crystals under study are given in the table. A large number of experimental structure amplitudes for each refined parameter and low ~1% R factors are evidence for a high reliability of the results.

The main and basically new result of this work is that a part of the Zn ions were located in the Nb positions at a concentration of 7.6% (table), i.e., at a concentration falling within the range of "threshold" changes in crystal properties [2]. Preliminary measurements carried out for crystals with 8.2% Zn lead to the same conclusion about the Zn location. At lower concentrations, Zn occupies only the Li positions, displacing Nb_{Li} and partially substituting for Li. This confirms the *a priori* suggestion, made on the basis of the available spectroscopic data [1, 2] and model calculations [8], about the interconnection between the threshold concentrations of non-photorefractive ions and their simultaneous insertion into two basis cationic positions.

Let us consider in more detail a "stepwise" transformation of the intrinsic defect structure of $LiNbO_3$ with increasing concentration of the inserted Zn ions. The discussion will be based on the model [4, 7]. Notice that the composition of the undoped crystal is consistent with the literature data for congruent crystals [3–5] and can be schematically represented by the well-known formula $[Li_{1-5x}Nb_x\square_{4x}][Nb]O_3$ ($\square$ stands for the vacancies, and the first and second square brackets denote the Li and Nb sites, respectively). In accordance with the above-mentioned *a priori* assumptions, the insertion of Zn into the Li position is accompanied by a decrease in the Nb_{Li} concentration (table); Nb_{Li} defects are fully

Selected crystallographic data for $LiNbO_3 : Zn_x$ single crystals; space group $R3c$

Atoms	site multiplicity	site occupancy	x/a	y/b	z/c	B_{eff}, Å
Zn: 0.0 mol %; $a = 5.149(1)$ Å; $c = 13.862(2)$ Å; $R = 0.011$, $R_w = 0.011$						
Nb	6	100.0(2)	0	0	0	0.41
(Li, Nb)	6	95.2(3)	0	0	0.2810(1)	0.99
O	18	100.0(2)	0.0482(1)	0.3432(1)	0.0645(2)	0.54
Zn: 2.87 mol %; $a = 5.150(1)$ Å; $c = 13.861(2)$ Å; $R = 0.009$, $R_w = 0.008$						
Nb	6	100.0(3)	0	0	0	0.39
(Li, Zn, Nb)	6	94.4(6)	0	0	0.27786(1)	1.00
O	18	100.0(5)	0.0479(1)	0.3429(1)	0.0640(2)	0.59
Zn: 5.2 mol %; $a = 5.151(1)$ Å; $c = 13.862(2)$ Å; $R = 0.011$, $R_w = 0.010$						
Nb	6	100.0(2)	0	0	0	0.45
(Li, Zn)	6	95.2(3)	0	0	0.2810(1)	0.95
O	18	100.0(2)	0.0482(1)	0.3432(1)	0.0645(2)	0.63
Zn: 7.6 mol %; $a = 5.152(1)$ Å; $c = 13.867(2)$ Å; $R = 0.016$, $R_w = 0.014$						
(Nb, Zn)	6	99.6(8)	0	0	0	0.42
(Li, Zn)	6	99.9(4)	0	0	0.2774(1)	0.89
O	18	100.0(5)	0.0485(1)	0.3438(1)	0.0633(2)	0.57

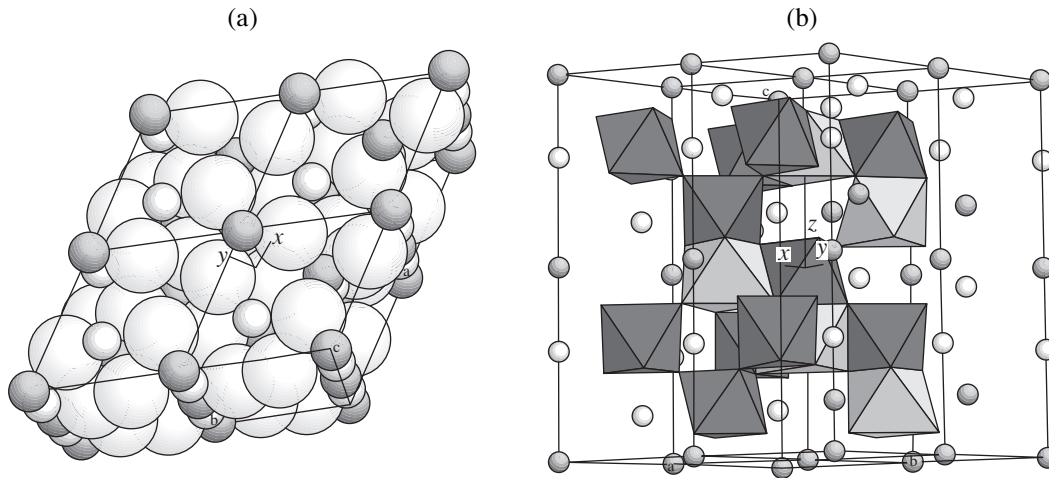(a)   (b)

Fig. 1. Crystal structure of lithium niobate $LiNbO_3$ of stoichiometric composition. (a) Close packing of the O atoms (large spheres) and positioning of the Nb atoms (dark spheres) and Li (light spheres) in octahedral cavities of the structure. (b) Nb and Li octahedra; a pair of such face-sharing octahedra is seen.

absent in a crystal with 5.2% Zn. According to the model [4, 6], the intrinsic Nb_{Li} defect is compensated by $4V_{Li}$; it is logical to assume that the Zn ion in the Li position, $Zn_{Li}^{\cdot}$, is compensated by V'_{Li} and formulate a crystal with 5.5% Zn as $[Li_{1-2x}Zn_x\square_x][[Nb]O_3$. Therefore, the replacement of Nb_{Li} by the Zn ions is simultaneously accompanied by a decrease in the concentration of Li vacancies. Such a transformation of the defect structure is corroborated by the fact that, at concentrations of 2.87 and 5.2% Zn, the lattice parameters *a* and *c* virtually coincide with their values in the undoped crystal (table), despite the difference in the ionic radii of Li^+ and Zn^{2-} [64 and 74 pm, respectively [14]). This departure from Vegard's law can be explained by the lattice "compaction" due to the collapsed vacancies [15].

The insertion of Zn into the Nb positions requires reformulation of the compensation conditions. Model calculation shows [8] that, if bivalent ions are inserted into both cationic positions, then the self-compensation mechanism, for which $Zn_{Nb}^{3'}$ is compensated by $3Zn_{Li}^{\cdot}$, is probable. The self-compensation requirement that the concentration ratio in a crystal with 7.6% Zn should be $[Zn_{Nb}]/[Zn_{Li}] = 1 : 3$ is fulfilled to a first approximation (table). This gives grounds to assume that the self-compensation mechanism is realized, and the crystal formula for $[Zn] \geq 7.6\%$ can be written as $[Li_{1-3y}Zn_{3y}][Nb_{1-y}Zn_y]O_3$. Since the self-compensation condition does not require that vacancies V_{Li} be involved, their concentration becomes zero at the threshold Zn concentration.

Figures 1a and 1b display the $LiNbO_3$ structure of stoichiometric composition. One can clearly imagine from these figures the changes occurring in the crystal structure upon doping it with zinc. The structure of

$LiNbO_3$ is a hexagonal close packing of O atoms, with two-thirds of octahedra each being occupied, respectively, by Nb and Li atoms while the last one-third being empty. Figure 1a demonstrates the close packing and the positioning of niobium and lithium in the octahedral cavities of the structure. For clarity, some of the Nb and Li octahedra are set off in Fig. 1b. If some of the Li positions in the matrix crystal are occupied by niobium, pairs of face-sharing Nb octahedra appear. Such uncommon configurations with two closely spaced highly charged cations are known in crystal chemistry. In the Zn-doped $LiNbO_3$ crystal, the pairs of two Nb octahedra are replaced by pairs of Nb and Zn octahedra; the first structural threshold appears when the Nb–Nb pairs of octahedra fully disappear (this occurs at concentrations $3 < [Zn] < 5\%$). In this concentration range, the concentration dependences of electrooptical coefficients z_{13} and z_{33} exhibit nonmonotonic behavior [15]; the other properties do not show any anomalies. If zinc substitutes for Nb in basis positions, the pairs of Li and Zn octahedra appear, while the pairs of Zn and Nb octahedra are retained. The Zn–Zn pairs of octahedra may also appear with low probability.

In summary, this study has confirmed the *a priori* assumption about the displacement of excess Nb from the Li positions by the Zn atoms and established the structural origin of the threshold Zn concentration (~7%), which is manifested as singularities in the concentration dependences of optical properties. The threshold Zn concentration corresponds to a change in the cationic position of the Zn ion and the disappearance of the Li vacancies from the lattice. This conclusion can likely be extended to the whole family of nonphotorefractive impurities (Mg, Zn, In, and Sc) in $LiNbO_3$ because of the similarity in the properties of these compositions.

REFERENCES

1. O. F. Schirmer, O. Thiemann, and M. Woehlecke, J. Phys. Chem. Solids **52**, 185 (1991).

2. T. Volk, M. Wöhlecke, N. Rubinina, *et al.*, Ferroelectrics **183**, 291 (1996).

3. S. C. Abrahams and P. March, Acta Crystallogr., Sect. B: Struct. Sci. **42**, 61 (1986).

4. N. Iyi, K. Kitamura, F. Izumi, *et al.*, J. Solid State Chem. **101**, 340 (1992).

5. N. Zotov, H. Boysen, F. Frey, *et al.*, J. Solid State Chem. **101**, 340 (1992).

6. P. Lerner, C. Legras, and J. P. Dumas, J. Cryst. Growth **3/4**, 231 (1968).

7. B. C. Grabmayer and F. Otto, J. Cryst. Growth **79**, 682 (1986).

8. H. Donnerberg, S. M. Tomlinson, C. R. A. Catlow, and O. F. Schirmer, Phys. Rev. B **44**, 4877 (1991).

9. B. C. Grabmayer, W. Wersing, and W. Koestler, J. Cryst. Growth **110**, 339 (1991).

10. N. Iyi, K. Kitamura, Y. Yajima, *et al.*, J. Solid State Chem. **118**, 148 (1995).

11. Z. I. Ivanova, A. I. Kovrigin, G. V. Luchinskii, *et al.*, Kvantovaya Élektron. (Moscow) **7**, 1013 (1980).

12. M. N. Zucker, E. Perenthaler, W. F. Kuhs, *et al.*, J. Appl. Crystallogr. **16**, 358 (1983).

13. T. S. Chernaya, N. N. Bydanov, L. A. Muradyan, *et al.*, Kristallografiya **33**, 75 (1988) [Sov. Phys. Crystallogr. **33**, 40 (1988)].

14. R. D. Shannon and C. T. Prewitt, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. **B25**, 925 (1969).

15. F. Abdi, M. Aillerie, M. Fontana, *et al.*, Appl. Phys. B: Lasers Opt. **B68**, 795 (1999).

Translated by V. Sakun

Quantum Bit Commitment in a Channel with Noise

S. N. Molotkov* and S. S. Nazin

Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia
* e-mail: molotkov@issp.ac.ru
Received December 5, 2000

Under rather general assumptions about the properties of a quantum channel with noise, a quantum protocol providing implementation of secret bit commitment with the probability as close to unity as is wished is proposed. © *2001 MAIK "Nauka/Interperiodica".*

PACS numbers: 03.67.Hk; 03.67.Dd

Weisner has come up in [1] with an idea that quantum physics can ensure higher security in information transmission than classical physics does. Later, after publications [2, 3], the secret key distribution (quantum cryptography) was studied in a number of papers. Apart from the key distribution protocol, other cryptographic protocols, so-called bit commitment (BC) and coin tossing (CT) protocols [4, 5], are important for applications and are of interest themselves. The quantum variant of these protocols was proposed in [6].

The BC protocol is an exchange protocol allowing two participants A and B, who do not trust each other, to implement the following scheme. Participant A sends a portion of information about his secret bit to participant B ($b = 0$ or 1, the commitment stage), so that, knowing this portion, participant B cannot determine what participant A contemplates. However, this portion of information is such that participant A cannot change the value of his secret bit b at the revelation stage when participant B claims the remaining portion of information. The CT protocol is a scheme through which two distant participants, who do not trust one another, can realize an honest coin tossing procedure.

The classical variants of the protocols are based on the unproved computational complexity of some trap functions (functions whose reversion requires exponentially large computation time with a classical computer) [7, 8].

It was believed that quantum protocols based not on computational complexity but on the fundamental exclusions imposed by quantum mechanics are unconditionally secure [9]. However, it was demonstrated later that the nonrelativistic quantum BC protocol is insecure [10, 11]. Participant A can cheat successfully by using a so-called Einstein–Podolsky–Rosen (EPR) attack [12]. The proof of the possibility of EPR attack is essentially based on the result obtained in [13] for measurements in quantum ensembles of composite systems.

All the above-mentioned nonrelativistic quantum protocols are based only on the properties of quantum states in the Hilbert space and do not explicitly include the effects of the state propagation from one distant participant of the protocol to the other. In practice, information is transmitted from one participant to another through Minkowski space–time. Allowance made for this fact in an explicit form extends the possibilities of constructing quantum relativistic protocols [14] and noticeably simplifies the proof of their security [15]. Restrictions imposed by special relativity on the measurability of quantum states allow one to realize the secret BC and CT protocols in an ideal channel [16]. An EPR attack is inefficient in the relativistic case because a state extended in space–time cannot be instantaneously measured. Moreover, it is impossible to instantaneously separate even one of a pair of orthogonal states with certainty. Restrictions imposed by special relativity on the measurability of quantum states were discussed even by Landau and Peierls [17].

In this paper, we propose the relativistic quantum BC protocol in a quantum channel with noise. The notion of the protocols is intuitively simple. Participant A prepares (turns on a source) one of a pair of orthogonal states corresponding to 0 or 1, which, being formed, are sent to a communication channel with the highest possible speed (speed of light c, which is taken as $c = 1$ in what follows). Until the whole states are available for participant B, he cannot with certainty determine what the value is of the secret bit, 0 or 1. Due to existing highest propagation speed, participant A cannot affect that portion of the state which has been sent to the communication channel; i.e., participant A is committed to the bit (commitment stage). When the whole state becomes available to participant B, he can determine the value of the secret bit with certainty, because of the orthogonality of the states, and can compare this value with the one that will be reported to him by participant A through a classical communication channel at the revelation stage. Restrictions imposed by special relativity on the measurability of quantum

states allow one to realize the original idea of the BC protocol about placing a portion of information at the disposal of one of the participants. The restriction of the accessible coordinate space automatically leads to the restriction of a part of the Hilbert space of states even for internal degrees of freedom of a quantum system (spin and polarization) because these degrees of freedom do not exist in isolation from the spatial degrees of freedom.

The protocol uses a pair of single-photon states which have orthogonal polarizations and spatial amplitude of a special form and correspond to 0 and 1:

$$|\psi_{0,1}\rangle = \int_0^\infty dk \mathcal{F}(k) a^+(k)|0\rangle \otimes |e_{0,1}\rangle$$

$$= \int_0^\infty dk \mathcal{F}(k)|k\rangle \otimes |e_{0,1}\rangle = |\mathcal{F}\rangle \otimes |e_{0,1}\rangle, \tag{1}$$

where $a^+(k)$ is the creation operator of the state with momentum (energy) $k > 0$, $\mathcal{F}(k)$ is the amplitude in the k representation, $|e_{0,1}\rangle$ is the polarization state, and

$$\int_0^\infty dk |\mathcal{F}(k)|^2 = 1, \quad [a(k), a^+(k')] = \delta(k - k'), \tag{2}$$

$$\langle e_i|e_j\rangle = \delta_{ij}, \quad i, j = 0, 1, \quad k \in (0, \infty).$$

In the space–time τ representation, the states have the form

$$|\psi_{0,1}\rangle = \int_{-\infty}^\infty d\tau \mathcal{F}(\tau)|\tau\rangle \otimes |e_{0,1}\rangle,$$

$$\mathcal{F}(\tau) = \int_0^\infty dk \mathcal{F}(k) e^{-ik\tau}, \quad \langle k|\tau\rangle = \frac{1}{\sqrt{2\pi}} e^{ik\tau}, \tag{3}$$

$$\tau = t - x, \quad \tau \in (-\infty, \infty),$$

where $\mathcal{F}(\tau)$ is the amplitude in the τ representation. This form corresponds to the intuitive notion of a packet that moves in the positive direction along the x axis with the speed of light and has a space–time form of $\mathcal{F}(\tau)$. The normalization condition in the τ representation, with allowance made for [18], has the form

$$\int_{-\infty}^\infty e^{ik\tau} \frac{1}{\tau + a} = i\pi \,\mathrm{sgn}(k) e^{-ika},$$

$$\langle\psi_{0,1}|\psi_{0,1}\rangle = \langle\mathcal{F}|\mathcal{F}\rangle = \int_{-\infty}^\infty\int_{-\infty}^\infty d\tau d\tau' \mathcal{F}(\tau)\mathcal{F}^*(\tau') \tag{4}$$

$$\times \left[\frac{1}{2}\delta(\tau - \tau') + \frac{i}{\pi}\frac{1}{\tau - \tau'}\right] = \int_{-\infty}^\infty |\mathcal{F}(\tau)|^2 d\tau.$$

The following two points are important for the protocol: (i) there is the highest speed of state propagation and (ii) the orthogonal states are indistinguishable with certainty if they are not available as a whole, even though they remain orthogonal under the restriction on the region accessible for measurements. The values of the classical bits 0 and 1 are attributed to two orthogonal polarization states $|e_0\rangle$ and $|e_1\rangle$. Because there is no polarization state in isolation from the spatial degrees of freedom $\mathcal{F}(\tau)$, for the states to be distinguished with certainty, i.e., with unit probability, the whole spatial region, where the amplitude $\mathcal{F}(\tau)$ is nonzero, should be accessible. More precisely, any measurement in the restricted τ region has nonzero error in distinguishing between states. Any measurement is described as a decomposition of unity in the Hilbert space $\mathcal{H}$ [19–23] and, when the restricted region $\Delta(\tau)$ is accessible [$\bar{\Delta}(\tau)$ is the complement with respect to the whole space $\tau \in (-\infty, \infty)$], is given as the following decomposition of unity:

$$I = \int_{-\infty}^\infty d\tau |\tau\rangle\langle\tau| \otimes I_{C^2}$$

$$= \int_{\Delta(\tau)} d\tau |\tau\rangle\langle\tau| \otimes (\mathcal{P}_0 + \mathcal{P}_1) + \int_{\bar{\Delta}(\tau)} d\tau |\tau\rangle\langle\tau| \otimes I_{C^2}, \tag{5}$$

$$\mathcal{P}_{0,1} = |e_{0,1}\rangle\langle e_{0,1}|,$$

where $\mathcal{P}_{0,1}$ are the projectors onto the polarization states $|e_{0,1}\rangle$. If the measurement outcome is in the region $\Delta(\tau)$ accessible for measurement, the probabilities of the outcomes in two orthogonal channels $\mathcal{P}_0$ and $\mathcal{P}_1$ have the form

$$\mathrm{Tr}\{\rho(0,1)(I(\Delta(\tau)) \otimes \mathcal{P}_{0,1})\}$$

$$= \int_{\Delta(\tau)} d\tau |\mathcal{F}(\tau)|^2 = N(\Delta(\tau)), \tag{6}$$

$$\mathrm{Tr}\{\rho(0,1)(I(\Delta(\tau)) \otimes \mathcal{P}_{0,1})\} \equiv 0,$$

where $\rho(0,1) = |\psi_{0,1}\rangle\langle\psi_{0,1}|$ and $N(\Delta(\tau))$ is the fraction of the outcomes in the accessible region. In this case, the probability of error due to the channel orthogonality is $p_e(\Delta(\tau)) = 0$. However, if the outcome is not obtained in the region accessible for measurements, the probability of error is $p_e(\bar{\Delta}(\tau)) = 1/2$ and the fraction of these outcomes is

$$\mathrm{Tr}\{\rho(0,1)(I(\bar{\Delta}(\tau)) \otimes I_{C^2})\}$$

$$= \int_{\bar{\Delta}(\tau)} d\tau |\mathcal{F}(\tau)|^2 = N(\bar{\Delta}(\tau)). \tag{7}$$

The total error is

$$P_e = p_e(\Delta(\tau))N(\Delta(\tau)) + p_e(\bar{\Delta}(\tau))N(\bar{\Delta}(\tau))$$

$$= 0 \cdot N(\Delta(\tau)) + \frac{1}{2}N(\bar{\Delta}(\tau)) = \frac{1}{2}\int_{\bar{\Delta}(\tau)} d\tau |\mathcal{F}(\tau)|^2 \neq 0. \tag{8}$$

The protocol uses states with the space–time amplitude of a special form, namely, the states consisting of two highly localized halves separated by the interval τ_0:

$$\mathcal{F}(\tau) = \frac{1}{\sqrt{2}}[f(\tau) + f(\tau - \tau_0)],$$

$$\int_{-\Delta\tau}^{\Delta\tau} d\tau |f(\tau)|^2 = \int_{-\Delta\tau + \tau_0}^{\Delta\tau + \tau_0} d\tau |f(\tau - \tau_0)|^2 = 1 - \delta, \tag{9}$$

$$\delta \ll 1, \quad \Delta\tau \ll \tau_0,$$

where δ is taken as small as is wished. The amplitude $f(\tau)$ cannot have a compact carrier [23]. However, the localization as strong as is wished and the falloff as close to exponential as is wished are possible [23, 24]. In what follows, we will omit the parameter δ, bearing in mind that it can be taken as the smallest parameter in the problem with any accuracy. The last statement means that, if the accessible region of the space–time τ covers the interval $-\Delta\tau < \tau < \Delta\tau + \tau_0$, the probability of error (8) is $P_e = 0$. If only one of the halves of the state is available, the probability of error (8) is $P_e = 1/4$. This means that, in order to distinguish between a pair of orthogonal states of form (9) with certainty, the space–time region of a size $\approx \tau_0$ should be accessible, which cannot be ensured in a time shorter than τ_0 because of the existence of the limiting propagation speed.

Participant A sends the initial states $\rho_{in}(0, 1) = (|e_{0, 1}\rangle \otimes |\mathcal{F}\rangle)(\langle\mathcal{F}| \otimes \langle e_{0, 1}|)$ to a quantum communication channel. The description of the quantum communication channel with noise reduces to the specification of an instrument [19–23], which is sometimes called superoperator, converting input density matrices into output ones which need not be normalized. Each quantum communication channel specifies an affine mapping of the set of input density matrices into the set of output density matrices. Any such mapping is reduced to the specification of the instrument $\mathcal{T}$:

$$\rho_{out}(0, 1) = \mathcal{T}[\rho_{in}(0, 1)]$$

$$= \int_{\infty}^{\infty}\int_{\infty}^{\infty} d\tau d\tau' \rho_{out}(\tau, \tau')|\tau\rangle\langle\tau'| \otimes \rho(e_0, e_1) \tag{10}$$

$$= \sum_{i=1}^{\infty} \lambda_i(|e_{i, 0, 1}\rangle \otimes |u_i\rangle)(\langle u_i| \otimes \langle e_{i, 0, 1}|),$$

where $|u_i\rangle = \int_{-\infty}^{\infty} d\tau u_i(\tau)|\tau\rangle$ are eigenvectors of the operator of an output density matrix, i.e., of the kernel $\rho_{out}(\tau, \tau')$. Taking Eq. (4) into account, we have

$$\int_{\infty}^{\infty} d\tau' \rho_{out}(\tau, \tau')u_i(\tau') = \lambda_i u_i(\tau),$$

$$\int_{\infty}^{\infty} d\tau u_i(\tau)u_j^*(\tau) = \delta_{ij}, \quad \sum_{i=1}^{\infty} \lambda_i \leq 1. \tag{11}$$

The output polarization vectors are $|e_{i, 0, 1}\rangle = \alpha_{i, 0, 1}|e_0\rangle + \beta_{i, 0, 1}|e_1\rangle$ ($|\alpha_{i, 0, 1}|^2 + |\beta_{i, 0, 1}|^2 = 1$). Any instrument can be represented in the form $\mathcal{T}[\rho] = \sum_i V_i \rho V_i^+$, with $\sum_i V_i V_i^+ \leq I$ [19–22] (here, it is sufficient to restrict our consideration to the discrete outcome space i). For the case under consideration, this representation can be written as

$$\mathcal{T}[\ldots] = \sum_{i=1}^{\infty} \lambda_i(|e_{i, 0, 1}\rangle \otimes |u_i\rangle)(\langle e_{0, 1}| \otimes \langle\mathcal{F}|)[\ldots]$$

$$\times (|\mathcal{F}\rangle \otimes |e_{0, 1}\rangle)(\langle u_i| \otimes \langle e_{i, 0, 1}|) + \mathcal{T}_\perp[\ldots], \tag{12}$$

where $\mathcal{T}_\perp[\ldots]$ is the instrument's part that identically gives zero on the linear shell of vectors $|\mathcal{F}\rangle \otimes |e_{0, 1}\rangle$.

Since the time of preserving the secret bit is determined by the extent τ_0 of the states, the channel can have an arbitrary length. Therefore, this length may be taken as 0 without loss of generality. In fact, the specification of an instrument is the description of the properties of the quantum communication channel, similarly to the classical case when the probability distribution over the output alphabet is specified for each symbol of the input alphabet. This mapping can intuitively be treated (with minor reservations) as the transformation of the input state $|\psi_{0, 1}\rangle$ with the form $\mathcal{F}(\tau)$ and polarization $e_{0, 1}$ to one of the output states, which have the form $u_i(\tau)$ and polarization $e_{i, 0, 1}$ and occur with the probability λ_i. In the case under consideration, the fact that the sum of probabilities does not exceed unity, i.e., $\sum_i \lambda_i \leq 1$, can be treated as disappearance (absorption) of a photon in the channel. The channel properties are specified by the functions $u_i(\tau)$ and probabilities λ_i, which are assumed to be *a priori* known and can be determined from the channel calibration. If one can choose a new interval $D\tau$ of the localization of the

halves of the output states so that

$$\forall i = 1, \infty, \quad \frac{1}{2} \int_{-D\tau}^{D\tau} d\tau |u_i(\tau)|^2 = \frac{1}{2} - \delta,$$

$$\frac{1}{2} \int_{-D\tau + \tau_0}^{D\tau + \tau_0} d\tau |u_i(\tau)|^2 = \frac{1}{2} - \delta, \quad D\tau \ll \tau_0, \tag{13}$$

where δ is as small as is wished, as in Eq. (9), the channel fits for implementing the proposed protocol. In other words, the channel is organized so that the states strongly localized at the input remain strongly localized at the output within an accuracy of $D\tau \ll \tau_0$ and $D\tau > \Delta\tau$ (see figure) but can change their form and polarization. The $D\tau$ value controls the accuracy with which a possible delay in the choice of the secret bit by participant A, i.e., the delay in sending the state to the communication channel, is detected. The probability of detecting the states at the output by participant B in the space–time window $\Delta(\tau)$ covering only one of the halves $u_i(\tau)$ independently of the outcome in the channels $\mathcal{P}_{0,1}$ is

$$\Pr\{\Delta(\tau)\} = \mathrm{Tr}\{\mathcal{T}[\rho_{in}(0,1)](I(\Delta(\tau)) \otimes I_{C^2})\}$$

$$= \sum_{i=1}^{\infty} \lambda_i \int_{\Delta(\tau)} d\tau |u_i(\tau)|^2 \leq \left(\frac{1}{2} - \delta\right) \sum_{i=1}^{\infty} \lambda_i \leq \frac{1}{2} - \delta \leq \frac{1}{2} \tag{14}$$

and can be made as close to 1/2 as is wished (with exponential accuracy by appropriately choosing $D\tau$ and τ_0). In this case, when only one half is available, i.e., during the time $\approx \tau_0$, the probability of distinguishing between the states is not higher than $1/2 \cdot 1/2 = 1/4$, according to Eq. (8).
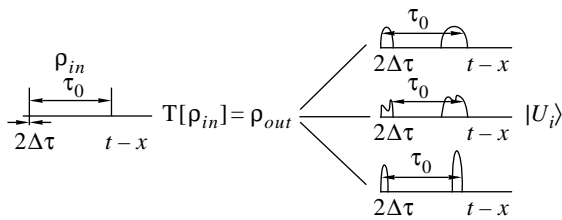
Let us now calculate the probability of error when the states become entirely available after the protocol time $D\tau + \tau_0 \approx \tau_0$ (recall that this distinguishing error is equal to zero in an ideal channel). If a state is entirely available (the time $\approx \tau_0$ after the protocol beginning), the probability of the outcome in one of the channels $\mathcal{P}_{0,1}$ is

$$\Pr\{\Delta(\tau) + \bar{\Delta}(\tau)\}$$

$$= \mathrm{Tr}\{\mathcal{T}[\rho_{in}(0,1)](I(\Delta(\tau)) \otimes I_{C^2})\} = \sum_{i=1}^{\infty} \lambda_i \leq 1. \tag{15}$$

The inequality $\Pr\{\Delta(\tau) + \bar{\Delta}(\tau)\} \leq 1$ means that the states do not all attain the channel output; i.e., states are absorbed in the channel with a probability of $1 - \sum_{i=1}^{\infty} \lambda_i$. Formally, with this probability they are never accessible for measurement by participant B. In this case, when the instrument of participant B does not change state, he can only guess what has been sent. The contribution of these events to the probability of error is $1/2(1 - \sum_{i=1}^{\infty} \lambda_i)$. Let us calculate the contribution to error from the events when the instrument of participant B changes state. The measurement minimizing error in distinguishing between two polarization states on the "honest" input states, which are sent by participant A, is described by the following decomposition of unity (see, e.g., [26] for more detail):

$$\sum_{i=1}^{\infty} \mathcal{P}_i \otimes (E_0 + E_1) + \mathcal{P}_\perp \otimes I_{C^2} = I \otimes I_{C^2}, \tag{16}$$

$$\mathcal{P}_i = |u_i\rangle\langle u_i|, \quad \mathcal{P}_\perp = I - \sum_{i=1}^{\infty} \mathcal{P}_i.$$

Here,

$$E_0 + E_1 = I_{C^2}, \quad E_0 = |\tilde{e}_0\rangle\langle\tilde{e}_0|,$$

$$I_{C^2} = |e_0\rangle\langle e_0| + |e_1\rangle\langle e_1|, \tag{17}$$

where $|\tilde{e}_0\rangle$ is the eigenvectors of the operator

$$\Gamma = \gamma_{00}|e_0\rangle\langle e_0| + \gamma_{01}|e_0\rangle\langle e_1| + \gamma_{10}|e_1\rangle\langle e_0| + \gamma_{11}|e_1\rangle\langle e_1|, \tag{18}$$

with

$$\gamma_{00} = \frac{1}{2}\sum_{i=1}^{\infty} \lambda_i(|\alpha_{i,1}|^2 - |\alpha_{i,0}|^2),$$

$$\gamma_{11} = \frac{1}{2}\sum_{i=1}^{\infty} \lambda_i(|\beta_{i,1}|^2 - |\beta_{i,0}|^2), \tag{19}$$

$$\gamma_{01} = \frac{1}{2}\sum_{i=1}^{\infty} \lambda_i(\alpha_{i,1}\beta_{i,0}^* - \alpha_{i,0}\beta_{i,1}^*), \quad \gamma_{10} = \gamma_{01}^*.$$

Taking into account Eq. (18) and the fact that *a priori* probabilities of choosing the states 0 or 1 by participant A are equal to 1/2, one can represent the total error in distinguishing between the polarization states on the

honest input states, when they are entirely available, in the form

$$P_e = \frac{1}{2}\left(1 - \sum_{i=1}^{\infty}\lambda_i\right) + \frac{1}{2}\text{Tr}\left\{\mathcal{T}[\rho_{in}(0)]\left(\left(\sum_{i=1}^{\infty}\mathcal{P}_i\right) \otimes E_1\right)\right\}$$
$$(20)$$
$$+ \frac{1}{2}\text{Tr}\left\{\mathcal{T}[\rho_{in}(1)]\left(\left(\sum_{i=1}^{\infty}\mathcal{P}_i\right) \otimes E_0\right)\right\} = \frac{1}{2} - |\gamma_2| < \frac{1}{2},$$

where γ_2 is the negative eigenvalue of the operator Γ appearing in Eq. (18):

$$\gamma_2 = \frac{1}{2}(\gamma_{00} + \gamma_{11}) - \frac{1}{2}\sqrt{(\gamma_{00} - \gamma_{11})^2 + 4|\gamma_{01}|^2}. \quad (21)$$

Assuming that the polarizations $|e_0\rangle$ and $|e_1\rangle$ are identically confused in the channel, one has $\gamma_2 = -|\gamma_{01}|$. As follows from Eqs. (18)–(20), the error in an ideal channel is $P_e = 0$.

The protocol is specified as follows. (i) The participants check only their local neighborhoods. Previously, they came to an agreement about the time of the protocol onset, the form $\mathcal{F}(\tau)$ of the states, and the polarization basis $|e_{0,1}\rangle$ for 0 and 1. (ii) Participant A encodes the secret bit b (0 or 1) by the parity bit over N states $\tilde{0}$ and $\tilde{1}$, consisting of the blocks of k states ($b = \sum_{j=1}^{N} \oplus a[i,j]$, $i = 1, \ldots, k$, where all $a[i,j]$ belonging to the same block are identical), and sends $k \cdot N$ states alternately through $k \cdot N$ quantum communication channels, while participant B carries out measurement (16). (iii) At the revelation stage at any $\Delta\tau < \tau < \Delta\tau + \tau_0$, participant B can request through the classical channel what has been sent by participant A. (iv) As the protocol time passes, participant B compares the outcomes of his measurements with what has been reported by participant A. (v) If all the tests are successful, the protocol is completed; otherwise it is broken.

So long as the protocol is not finished, the probability of the correct determination of the secret bit by participant B exceeds 1/2, which is the probability of simple guessing, only by an exponentially small quantity. Indeed, the number of binary rows of length $k \cdot N$ in the block representation with 0 and 1 is (see [27] for the summation)

$$N_{odd} = N_{even} = \frac{1}{2}\sum_{m=0}^{N} C_{N \cdot k}^{m \cdot k}$$
$$(22)$$
$$= \frac{2^{N \cdot k}}{2k}\sum_{l=1}^{k} \cos^{N \cdot k}\left(\frac{l\pi}{k}\right)\cos(lN\pi) \approx \frac{1}{2K}2^{N \cdot k}$$

and is virtually equal to the total number of binary rows of length $N \cdot k$. Shannon's information [28–30] in the set of the block rows is (within an accuracy of round-

off error) the number of binary symbols that are necessary for identifying the row parity:

$$I = \log_2\left(\frac{2^{N \cdot k}}{2k}\sum_{l=1}^{k} \cos^{N \cdot k}\left(\frac{l\pi}{k}\right)\cos(lN\pi)\right) \approx \eta N \cdot k,$$
$$(23)$$
$$\eta \approx 1;$$

i.e., almost all bits in the row should be known. However, when only half of the state is accessible ($\Delta\tau < \tau < \Delta\tau + \tau_0$), the probability of error in determining a bit in each position is equal to or less than 1/4, according to Eq. (8), even in a channel without noise. Therefore, the probability that participant B has determined the parity bit until the protocol is finished does not exceed

$$P(\text{parity}) = \frac{1}{2} + 2^{-\frac{\eta}{2}N \cdot k}. \quad (24)$$

Let us calculate the probability of error in determining the parity bit by participant B after the protocol is finished. The block representation for even k is stable (errors are corrected by the majority voting method) when the number of errors in each block is no more than $k/2 - 1$. The probability of error in the block $\tilde{0}$ or $\tilde{1}$ is

$$P_e(k) = \sum_{i=\frac{k}{2}}^{k} C_k^i P_e^i (1 - P_e)^{k-i}$$
$$(25)$$
$$\approx \sqrt{\frac{2}{\pi k}}[2\sqrt{P_e(1 - P_e)}]^k,$$

which may be made smaller than any preset quantity by choosing k. The total error in determining the parity bit is (N is assumed to be even)

$$P_e(\text{parity}) = \sum_{i=odd}^{N-1} C_N^i P_e^i(k)(1 - P_e(k))^{N-i}, \quad (26)$$

where the summation is performed only over odd superscripts i because error in the parity bit arises when malfunction occurs in an odd number of blocks. Taking into account that

$$\frac{1}{2}[(x + y)^N - (x - y)^N] = \sum_{i=odd}^{N-1} C_N^i x^i y^{N-i} \quad (27)$$

and setting $z = P_e(k)$ and $y = 1 - P_e(k)$ ($x + y = 1$), we find

$$P_e(\text{parity}) = \frac{1}{2}[1 - (1 - 2P_e(k))^N]. \quad (28)$$

By choosing large enough k for a given quantum channel, the probability $P_e(k)$ can be made as small as is wished so that $NP_e(k) \ll 1$ is exponentially small. In this case, the error in determining the parity bit after the

protocol is finished is also as small as is wished and the probability of correct determination is as close to unity as is wished.

Let us discuss now the stability of the protocol to deception by participant A. Because the minimal distance (according to Hamming) between two block rows of different parities equals k (the minimal number of different positions), k positions, as a minimum, should be changed for changing the parity of the whole row. Since the probability of correct determination of each block $\tilde{0}$ or $\tilde{1}$ is no less than $1 - P_e(k) \longrightarrow 1$ [$P_e(k)$ is exponentially small, see Eq. (25)], the probability that deception by participant A cannot be detected by participant B does not exceed this quantity.

The protocol is also stable to the delay in choosing the secret bit by participant A after the protocol onset. Recall that, for honest undelayed input states, the probability of outcome in the channel $\mathscr{P}_\perp = I - \sum_{i=1}^{\infty} \mathscr{P}_i$ is equal to zero:

$$\Pr(\Delta(\tau) + \bar{\Delta}(\tau)) = \mathrm{Tr}\{(\mathscr{T}[\rho_{\mathrm{in}}(0,1)] + \mathscr{T}_\perp[\rho_{\mathrm{in}}(0,1)])$$

$$\times (\mathscr{P}_\perp \otimes I_{C^2})\} = \mathrm{Tr}\left\{\left(\left(\sum_i \lambda_i |\alpha_{i,0,1}|^2 \mathscr{P}_i\right) \otimes |e_0\rangle\langle e_0|\right.\right.$$

$$+ \left(\sum_i \lambda_i \alpha_{i,0,1} \beta_{i,0,1}^* \mathscr{P}_i\right) \otimes |e_0\rangle\langle e_1|$$

$$+ \left(\sum_i \lambda_i \beta_{i,0,1} \alpha_{i,0,1}^* \mathscr{P}_i\right) \otimes |e_1\rangle\langle e_0| \qquad (29)$$

$$+ \left.\left(\sum_i \lambda_i |\beta_{i,0,1}|^2 \mathscr{P}_i\right) \otimes |e_1\rangle\langle e_1|\right)$$

$$\times \left(I - \sum_j \mathscr{P}_j\right) \otimes I_{C^2}\right\} = 0,$$

with allowance made for $|\alpha_{i,0,1}|^2 + |\beta_{i,0,1}|^2 = 1$ and $\mathscr{P}_i\mathscr{P}_j = \delta_{ij}\mathscr{P}_i$.

Any state delayed by more than $D\tau$ can be detected with the probability as close to unity as is wished. This detection can be ensured by the restrictions imposed on instrument (12) by special relativity, more precisely, by existence of a limiting propagation speed. If a state prepared at the input of any quantum communication channel is strongly localized, i.e., the squared amplitude is $\int_{-\Delta\tau}^{\Delta\tau} d\tau |\mu(\tau)|^2 = 1 - \delta$, where δ is exponentially small and $\Delta\tau \longrightarrow 0$, this state cannot be detected at the channel output in a time shorter than $t = L/c$, where L is the channel length; more precisely, detection with the probability as close to unity as is wished is possible in the time window $-\Delta\tau + L/c \leq t \leq \Delta\tau + L/c$. For the case

under consideration, instrument (12) must convert states prepared later at the input into states which appear later at the output. The shift value for the leading edge of the state amplitude at the output cannot be less than the corresponding shift at the input.

Any delayed input state can be presented in the form (polarization degrees of freedom are omitted for brevity)

$$\rho_{\mathrm{delay}} = \sum_l \mu_l |\mu_l\rangle\langle\mu_l|, \quad \sum_l \mu_l = 1,$$

$$|\mu_l\rangle = \int_{-\infty}^{\infty} d\tau \mu_l(\tau)|\tau\rangle, \qquad (30)$$

where $|\mu_l\rangle$ are the eigenstates of the density matrix and the functions $\mu_l(\tau)$ have a carrier which does not overlap in the interval $D\tau$ with the carrier of the leading half of any of the functions $u_i(\tau)$ obtained at the channel output from undelayed states. At the channel output, the density matrix ρ_{delay} is converted to such a density matrix that the carriers of the eigenfunctions $|\eta_k\rangle$ also do not overlap with the leading half $u_i(\tau)$ in the interval $D\tau$:

$$(\mathscr{T} + \mathscr{T}_\perp)[\rho_{\mathrm{delay}}] = \sum_k \eta_k |\eta_k\rangle\langle\eta_k|,$$

$$\sum_k \eta_k \leq 1, \quad |\eta_k\rangle = \int_{-\infty}^{\infty} d\tau \eta_k(\tau)|\tau\rangle. \qquad (31)$$

Therefore, $|\langle\eta_k|u_i\rangle|^2 \leq 1/2$, because $\eta_k(\tau)$ does not cover the leading half $u_i(\tau)$ where one-half of the squared modulus of $u_i(\tau)$ is gained.

The probability of the outcome on delayed states in the channel $\sum_i \mathscr{P}_i < I$ is

$$\mathrm{Tr}\left\{\left(\sum_k \eta_k |\eta_k\rangle\langle\eta_k|\right)\left(\sum_i \mathscr{P}_i\right)\right\} < 1, \qquad (32)$$

and the probability of the outcome on undelayed states is equal to unity. Similarly, the probability of the outcome in the channel $\mathscr{P}_\perp = I - \sum_{i=1}^{\infty} \mathscr{P}_i$ is (recall that this probability is equal to zero on honest states)

$$\mathrm{Tr}\left\{\left(\sum_k \eta_k |\eta_k\rangle\langle\eta_k|\right)\left(I - \sum_i \mathscr{P}_i\right)\right\} = p_\perp \neq 0. \qquad (33)$$

The sum of probabilities of the outcomes in both channels is equal to unity if all states attain the channel output, i.e., are not absorbed, $\sum_k \eta_k = 1$.

Delay in choosing the secret bit (state delay) is detected upon appearance of the outcomes in the channel $\mathscr{P}_\perp$ with the probability $p_\perp$. For changing parity bit,

it is sufficient to delay the states only in one of the blocks of k states. The probability that delay of k states by participant A remains undetected is equal to the probability that all k delayed states do not give an outcome in the channel $\mathcal{P}_\perp$ and, therefore, imitate the statistics of measurement outcomes for honest states. This means that

$$P_{\text{cheat}} = (1 - p_\perp)^k \ll 1, \tag{34}$$

which can be achieved by taking large enough k for known probability $p_\perp$.

Thus, the protocol provides implementation of honest bit commitment protocol with the probability as close to unity as is wished.

REFERENCES

1. S. Wiesner, SIGACT News **15**, 78 (1983).
2. A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
3. C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992); C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).
4. M. Blum, in *Proceedings of the 24th IEEE Computer Conference, 1982*, p. 133; SIGACT News **15**, 23 (1983).
5. P. A. Feldman, in *Proceedings of the 28th Annual Symposium on Foundations on Computer Science, 1987*, p. 469.
6. C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computer Systems and Signal Processing, Bangalore, 1984*, p. 175.
7. P. W. Shor, in *Proceedings of the 35th Annual Symposium on Foundations on Computer Science, Santa Fe*, Ed. by S. Goldwasser (IEEE Computer Society Press, Los Alamitos, 1994), p. 124.
8. A. Yu. Kitaev, Usp. Mat. Nauk **52** (6), 54 (1997).
9. C. Cré peau, What is Going on with Quantum Bit Commitment, Crypto-96.
10. H.-K. Lo and H. F. Chau, Phys. Rev. Lett. **78**, 3410 (1997).
11. D. Mayers, Phys. Rev. Lett. **78**, 3414 (1997).
12. A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
13. L. P. Hugston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).
14. L. Goldenberg and L. Vaidman, Phys. Rev. Lett. **75**, 1239 (1995); quant-ph/9506030.
15. S. N. Molotkov and S. S. Nazin, quant-ph/0008008.
16. S. N. Molotkov and S. S. Nazin, quant-ph/9911055; quant-ph/9910034; Zh. Éksp. Teor. Fiz. **117**, 818 (2000) [JETP **90**, 714 (2000)]; Pis'ma Zh. Éksp. Teor. Fiz. **70**, 684 (1999) [JETP Lett. **70**, 702 (1999)].
17. L. D. Landau and R. Peierls, Z. Phys. **69**, 56 (1931); *Collection of Works* (Nauka, Moscow, 1969), Vol. 1, p. 56.
18. Yu. A. Brychkov and A. P. Prudnikov, *Integral Transformations of Generalized Functions* (Nauka, Moscow, 1977).
19. E. B. Devis, *Quantum Theory of Open Systems* (Academic, London, 1976).
20. A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982); A. S. Holevo, *Lectures on Statistical Structure of Quantum Theory* (1999), pp. 1–177.
21. K. Kraus, *States, Effects and Operations* (Springer-Verlag, Berlin, 1983).
22. P. Busch, M. Grabowski, and P. J. Lahti, in *Springer Lecture Notes in Physics*, Vol. 31: *Operational Quantum Physics* (Springer-Verlag, Berlin, 1995).
23. N. N. Bogolyubov, A. A. Logunov, and I. T. Todorov, *Introduction to Axiomatic Quantum Field Theory* (Nauka, Moscow, 1969; Benjamin, New York, 1975).
24. I. Bialynicki-Birula, Phys. Rev. Lett. **80**, 5247 (1998).
25. W. F. Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955).
26. C. A. Fuchs, quant-ph/9601020.
27. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1981; Gordon and Breach, New York, 1986).
28. C. E. Shannon, Bell Syst. Tech. J. **27**, 397 (1948); **27**, 623 (1948).
29. R. G. Gallager, *Information Theory and Reliable Communication* (Wiley, New York, 1968; Sov. Radio, Moscow, 1974).
30. I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discreter Memoryless Systems* (Akadémiai Kiadó, Budapest, 1981).

Translated by R. Tyapaev

Erratum: "Search for T-Odd Left–Right Asymmetry of Prompt Neutron Emission in Binary Fission of the ^{233}U and ^{239}Pu Nuclei by Slow Polarized Neutrons," [*JETP Letters*, 72, no. 12, 593–594 (2000)]

G. V. Danilyan, V. A. Krakhotin, V. S. Pavlov, A. V. Fedorov, E. I. Korobkina, and E. Lelivre-Berna

In this paper, the authors missed one sentence at the end of the main text: "This work was supported by the Russian Foundation for Basic Research (project no. 00-02-16011) and the program "Fundamental Nuclear Physics" (project no. 134-06)."

On the Energy Loss of High Energy Quarks in a Finite-Size Quark–Gluon Plasma[1]

B. G. Zakharov

Institut für Kernphysik, Forschungszentrum Jülich, D-52425 Jülich, Germany
Landau Institute for Theoretical Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia
Received December 25, 2000

The induced gluon emission from a fast quark passing through a finite-size QCD plasma is studied within the light-cone path integral approach. It is shown that the leading log approximation used in previous studies fails when the gluon formation length becomes on the order of the length of the medium traversed by the quark. Calculation of the energy loss beyond the leading log approximation gives the energy loss which grows logarithmically with quark energy, contrary to the energy-independent prediction of the leading log approximation. © 2001 MAIK "Nauka/Interperiodica".

PACS number: 12.38.Mh

In recent years, much attention has been attracted to the problem of induced gluon radiation from fast partons in a hot QCD medium (for a review, see [1]). It is of great interest in connection with the current experiments at SPS, RHIC, and future experiments at LHC on $A + A$ collisions, since jet quenching due to the parton energy loss can be a good probe for the formation of a hot quark–gluon plasma (QGP).

Evaluation of the gluon emission from a fast parton in a medium requires the understanding of the nonabelian analogue of the Landau–Pomeranchuk–Migdal (LPM) effect [2, 3]. There are two approaches to the LPM effect in QCD: the so-called BDMS approach [4] (see also [1, 5]) based on the Feynman diagrammatic formalism, and the light-cone path integral (LCPI) approach developed in our paper [6] (see also [7–10]). The BDMS approach neglects the mass effects and applies in the case of large suppression of the radiation rate, as compared to the Bethe–Heitler one. The LCPI approach applies to an arbitrary strength of suppression. For large suppression, these approaches are equivalent [1, 4, 11]. The probability of gluon emission in the BDMS and LCPI approaches is expressed through the solution of a two-dimensional Schrödinger equation with an imaginary potential. This equation describes evolution of the color singlet $\bar{q}gq$ system in a medium. The potential is proportional to the cross section for scattering of the $\bar{q}gq$ system by a medium constituent. For the QGP, the constituents can be modeled as Debye-screened colored Coulomb scattering centers [12].

In [4], the quark energy loss ΔE has been evaluated analytically by treating interaction of the $\bar{q}gq$ system with the Debye-screened centers in the Leading Log Approximation (LLA), which is equivalent to the harmonic oscillator approximation for the Hamiltonian of the $\bar{q}gq$ system. For a quark produced inside a finite-size QGP, the BDMS prediction is

$$\Delta E_{\text{BDMS}} = \frac{C_F \alpha_s}{4} \frac{L^2 \mu^2}{\lambda_g} \tilde{v},\qquad(1)$$

where L is the length of QGP traversed by the quark, μ is the Debye screening mass, λ_g is the mean free path of the gluon in QGP, C_F is the color Casimir for the quark, and the factor $\tilde{v}$ grows smoothly with L; at $L \gg \lambda_g$, $\tilde{v} \approx \log(L/\lambda_g)$.

The energy-independent ΔE, Eq. (1), differs from that obtained recently by Gyulassy, Levai, and Vitev (GLV) [13]. By calculating the Feynman diagrams for single scattering [first order ($N = 1$) in opacity], they obtained

$$\Delta E_{\text{GLV}} = \frac{C_F \alpha_s}{4} \frac{L^2 \mu^2}{\lambda_g} \log\frac{E}{\mu}.\qquad(2)$$

Since ΔE_{BDMS} should include the $N = 1$ contribution, the contradiction between Eqs. (1) and (2) at $E \longrightarrow \infty$ seems to be surprising.[2] Until now, there has not been given any explanation of this fact, except for the argument of the authors of [13] that it can be connected with the neglect of the finite kinematic bounds in the analysis [4]. However, it is clear that it cannot be important at $E \longrightarrow \infty$.

---

[1] This article was submitted by the author in English.

[2] Strictly speaking, the derivation of the BDMS formalism given in [4] is only valid when the number of rescatterings is large. However, since the formulas obtained are equivalent to those of the LCPI [6] approach, which is free from this restriction, it is clear that the BDMS prediction should contain the $N = 1$ term.

In this paper, the above puzzle of discrepancy between the BDMS and GLV predictions is resolved. It is demonstrated that the absence of the logarithmic energy dependence in Eq. (1) is connected with the fact that the LLA fails when the gluon formation length becomes on the order of L. In this case, the spectrum is dominated by the $N = 1$ scattering, which simply vanishes in the LLA. It is shown that, if one uses the actual imaginary potential, the energy loss grows logarithmically with quark energy. However, the denominator in the argument of the logarithm is not the Debye mass, as it is in Eq. (2).

We will work in the LCPI formalism [6]. The probability distribution of the induced gluon emission from a quark produced at $z = 0$ can be written as [10]

$$\frac{dP}{dx} = \int_0^\infty dz\, n(z) \frac{d\sigma_{\text{eff}}^{BH}(x, z)}{dx}, \tag{3}$$

where x is the gluon fractional momentum, n is the number density of the medium, and

$$\frac{d\sigma_{\text{eff}}^{BH}(x, z)}{dx} = \text{Re} \int d\boldsymbol{\rho}\, \Psi^*(\boldsymbol{\rho}, x) \sigma_3(\rho, x) \Psi_m(\boldsymbol{\rho}, x, z). \tag{4}$$

Here, σ_3 is the cross section for interaction of the $\bar{q}gq$ system with a scattering center. The relative transverse separations in the $\bar{q}gq$ system are $\boldsymbol{\rho}_{g\bar{q}} = (1 - x)\boldsymbol{\rho}$ and $\boldsymbol{\rho}_{q\bar{q}} = -x\boldsymbol{\rho}$, $\Psi(\boldsymbol{\rho}, x)$ is the light-cone wave function for the $q \longrightarrow gq$ transition in vacuum, and $\Psi_m(\boldsymbol{\rho}, x, z)$ is the quark light-cone wave function in the medium for the longitudinal coordinate z (spin and color indices are omitted). The wave functions (modulo color factor) read

$$\Psi(\boldsymbol{\rho}, x) = P(x)\left(\frac{\partial}{\partial \rho'_x} - is_g \frac{\partial}{\partial \rho'_y}\right)$$

$$\times \int_0^\infty d\xi \exp\left(-\frac{i\xi}{L_f}\right) \mathcal{K}_0(\boldsymbol{\rho}, \xi | \boldsymbol{\rho}', 0)\Bigg|_{\boldsymbol{\rho}' = 0}, \tag{5}$$

$$\Psi_m(\boldsymbol{\rho}, x, z) = P(x)\left(\frac{\partial}{\partial \rho'_x} - is_g \frac{\partial}{\partial \rho'_y}\right)$$

$$\times \int_0^z d\xi \exp\left(-\frac{i\xi}{L_f}\right) \mathcal{K}(\boldsymbol{\rho}, z | \boldsymbol{\rho}', z - \xi)\Bigg|_{\boldsymbol{\rho}' = 0}, \tag{6}$$

where $P(x) = i\sqrt{\alpha_s/2x}\,[s_g(2 - x) + 2s_q x]/2M(x)$, $s_{q, g}$ denote the parton helicities, $\mathcal{K}$ is the Green's function for the two-dimensional Hamiltonian

$$\hat{H}(z) = -\frac{1}{2M(x)}\left(\frac{\partial}{\partial \boldsymbol{\rho}}\right)^2 - i\frac{n(z)\sigma_3(\rho, x)}{2}, \tag{7}$$

and

$$\mathcal{K}_0(\boldsymbol{\rho}_2, z_2 | \boldsymbol{\rho}_1, z_1)$$

$$= \frac{M(x)}{2\pi i(z_2 - z_1)} \exp\left[\frac{iM(x)(\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1)^2}{2(z_2 - z_1)}\right] \tag{8}$$

is the Green's function for Hamiltonian (7) with $v(\boldsymbol{\rho}, z) = 0$, $M(x) = Ex(1 - x)$, and $L_f = 2Ex(1 - x)/\epsilon^2$, where $\epsilon^2 = m_g^2(1 - x) + m_q^2 x$. The gluon mass m_g plays the role of infrared cutoff removing the contribution from long-wavelength gluons which cannot propagate in the QGP. It is natural to take $m_g \sim \mu$. However, for large suppression, which occurs at $E \longrightarrow \infty$, the parton masses can simply be neglected. The three-body cross section can be written as [14]

$$\sigma_3(\rho, x)$$

$$= \frac{C_A}{2C_F}\left[\sigma_2((1 - x)\rho) + \sigma_2(\rho) - \frac{1}{N_c^2}\sigma_2(x\rho)\right], \tag{9}$$

where $C_A = N_c$ is the octet color Casimir and $\sigma_2(\rho)$ is the dipole cross section for scattering of a $\bar{q}q$ pair by a color center. For the parametrization $\sigma_2(\rho) = C_2(\rho)\rho^2$, the factor C_2 is

$$C_2(\rho) = \frac{C_T C_F \alpha_s^2}{\rho^2} \int d\mathbf{q}\, \frac{[1 - \exp(i\mathbf{q}\boldsymbol{\rho})]}{(q^2 + \mu^2)^2}. \tag{10}$$

Here, C_T is the color Casimir of the scattering center. In the region $\rho \ll 1/\mu$, which dominates the spectrum for strong suppression, Eq. (10) takes the form

$$C_2(\rho) \approx \frac{C_F C_T \alpha_s^2 \pi}{2} \log\left(\frac{1}{\rho\mu}\right). \tag{11}$$

The LLA consists in replacing $C_2(\rho)$ by $C_2(\rho_{\text{eff}})$, where ρ_{eff} is the typical value of ρ. This seems to be a reasonable procedure, since $C_2(\rho)$ has only a slow logarithmic dependence on ρ. Then $\sigma_3(\rho, x) = C_3(x)\rho^2$, where $C_3(x) = C_2(\rho_{\text{eff}})A(x)$, with $A(x) = [1 + (1 - x)^2 - x^2/N_c^2]C_A/2C_F$, and Hamiltonian (7) takes the oscillator form with frequency $\Omega(x) = \sqrt{-iC_3(x)n/M(x)}$. The value of ρ_{eff} is connected with the gluon formation length l_f by the Schrödinger diffusion relation $\rho_{\text{eff}}^2 \sim l_f/2M$. l_f is the typical scale of ξ in Eqs. (5) and (6) when the wave functions are substituted in Eq. (4).

Let us discuss the gluon emission at a qualitative level. We begin with estimating ρ_{eff} and l_f. Let us first estimate these quantities for gluon emission from a quark in an infinite medium. We will denote them by $\bar{\rho}_{\text{eff}}$ and $\bar{l}_f$. They should also be related to each other by the Schrödinger diffusion relation. On the other hand, the absorption effects for the $\bar{q}gq$ system should

become strong at the scale $\bar{l}_f$. It means that $\bar{l}_f n C_3 \bar{\rho}_{\text{eff}}^2/2 \sim 1$. From these conditions one gets $\bar{\rho}_{\text{eff}} \sim [E_a x(1 - x)nC_3]^{-1/4}$ and $\bar{l}_{\text{eff}} \sim 2\sqrt{E_a x(1 - x)/nC_3}$. These estimates are valid when $\bar{\rho}_{\text{eff}} \lesssim 1/\epsilon$ and $\bar{l}_f \lesssim L_f$.

Let us now turn to the gluon emission from a quark produced inside a finite-size medium. In this case, two qualitatively different situations are possible in the high-energy limit. The first regime occurs for the gluons with x such that $\bar{l}_f \lesssim L$. In this case, the finite-size effects play a marginal role and $\rho_{\text{eff}} \sim \bar{\rho}_{\text{eff}}$. The spectrum can roughly be calculated using the effective Bethe–Heitler cross section for the infinite medium. We call this regime the infinite-medium regime. The second regime occurs for the gluons for which $\bar{l}_f \gtrsim L$. In this case, $\rho_{\text{eff}} \sim \rho_d(L)$, where $\rho_d(L) = \sqrt{L/2M}$ is simply the diffusion radius on the scale of quark path length inside the medium. In this regime, the effective Bethe–Heitler cross section is chiefly controlled by the finite-size effects. We will call this regime the diffusion regime. Thus, one can write for the above two regimes

$$\rho_{\text{eff}} \sim \min(\bar{\rho}_{\text{eff}}, \rho_d(L), 1/\epsilon). \quad (12)$$

Here, it is taken into account that $\rho_{\text{eff}} \lesssim 1/\epsilon$. In terms of x, the infinite-medium regime occurs at $x \lesssim \delta$ and $(1 - x) \lesssim \delta$ and the diffusion regime occurs at $\delta \lesssim x \lesssim (1 - \delta)$, where

$$\delta \sim nC_3 L^2/4E. \quad (13)$$

For definiteness, only the region $x \lesssim 0.5$ is discussed below. At $x \gtrsim \delta$, the probability of interaction of the $\bar{q}gq$ system with the medium [it is of the order of $n\sigma_3(\rho_d, x)L$] becomes small. Thus, it is clear that, in the developed diffusion regime, the spectrum is dominated by the $N = 1$ scattering. It is surprising that this proves to be in apparent contradiction with the prediction of the LLA. The LLA spectrum can be obtained by using the oscillator Green's function in Eq. (6). For zero parton masses, it gives

$$\frac{dP}{dx} = -\frac{2G(x)}{\pi}\text{Re}\int_0^L dz \int_0^z d\xi \frac{\Omega^2}{\cos^2\Omega\xi}$$
$$= \frac{2G(x)}{\pi}\ln|\cos\Omega L|, \quad (14)$$

where $G(x) = \alpha_s C_F[1 - x + x^2/2]/x$. This spectrum was derived in [4]. Note that $|\Omega L| \sim 1$ at $x \sim \delta$. For the diffusion regime, one gets from Eq. (14)

$$\frac{dP}{dx}\bigg|_{x \gg \delta}^{\text{LLA}} \approx \frac{G(x)C_3^2 n^2 L^4}{8\pi E^2 x^2(1 - x)^2}. \quad (15)$$

Since the right-hand side of Eq. (15) $\propto n^2$, it is clear that it corresponds to the $N = 2$ term. Thus, one sees that the $N = 1$ contribution is simply absent in the LLA.

The fact that the LLA fails in the diffusion regime can be directly seen from the calculation of the $N = 1$ contribution. To obtain it, one should use free Green's function (8) in Eq. (6). Then, in the massless limit, Eq. (4) gives

$$\frac{d\sigma_{\text{eff}}^{\text{BH}}(x, z)}{dx}\bigg|_{N = 1} = \frac{G(x)A(x)M(x)}{2\pi}$$
$$\times \text{Im}\int_0^z \frac{d\xi}{\xi^2}\int_0^\infty d\rho^2 \rho^2 C_2(\rho)\exp\left(\frac{iM(x)\rho^2}{2\xi}\right). \quad (16)$$

For $C_2(\rho) = \text{const}$, the ρ^2 integral in Eq. (16) has zero imaginary part and the right-hand side of Eq. (16) is also zero. On the other hand, using Eq. (11) one gets from Eq. (16)

$$\frac{d\sigma_{\text{eff}}^{\text{BH}}(x, z)}{dx}\bigg|_{N = 1} = \frac{\alpha_s^2 \pi C_T C_F G(x)A(x)z}{4Ex(1 - x)}. \quad (17)$$

Then Eq. (3) yields

$$\frac{dP}{dx}\bigg|_{N = 1} = \frac{\alpha_s^2 \pi C_T C_F G(x)A(x)nL^2}{8Ex(1 - x)}. \quad (18)$$

Let us see why the LLA fails in momentum representation, in which Eq. (4) reads

$$\frac{d\sigma_{\text{eff}}^{\text{BH}}(x, z)}{dx} = \frac{\alpha_s^2 C_T C_F A(x)}{(2\pi)^2}$$
$$\times \text{Re}\int d\mathbf{p}\,d\mathbf{q}\frac{[\Psi^*(\mathbf{p}, x) - \Psi^*(\mathbf{p} - \mathbf{q}, x)]\Psi_m(\mathbf{p}, x, z)}{(q^2 + \mu^2)^2}. \quad (19)$$

In the massless limit, one can obtain from Eq. (19)

$$\frac{d\sigma_{\text{eff}}^{\text{BH}}(x, z)}{dx}\bigg|_{N = 1}$$
$$= \frac{\alpha_s^2 C_T C_F G(x)A(x)}{2\pi}\int dp^2 dq^2 \frac{F(p, q)}{(q^2 + \mu^2)^2}, \quad (20)$$

$$F(p, q)$$
$$= \text{Re}\frac{1}{p^2}\left[1 - \exp\left(-\frac{izp^2}{2M(x)}\right)\right]\int_0^{2\pi} d\phi\frac{\mathbf{q}(\mathbf{q} - \mathbf{p})}{(\mathbf{q} - \mathbf{p})^2}, \quad (21)$$

where ϕ is the angle between $\mathbf{q}$ and $\mathbf{p}$. The logarithmic situation with dominance of $q^2 \ll p^2$ would correspond to $F(p, q) \propto q^2$ at $q^2 \ll p^2$. However, the azimuthal ϕ integral in Eq. (21) equals $2\pi\theta(q^2 - p^2)$, and the process is dominated by hard t-channel exchanges with $q^2 > p^2 \sim 2M(x)/z$. After integrating over p^2 and q^2 in Eq. (20), one reproduces Eq. (17).

It should be emphasized that the LLA fails only in the diffusion regime. But it is a good approximation in the infinite-medium regime when Ψ_m falls off rapidly at a scale much smaller than $\rho_d(L)$. It is also worth noting that boundary (13), beyond which the diffusion regime occurs, is obtained assuming that LPM suppression in the infinite-medium regime is strong [it means that $\bar{\rho}_{eff}(x \sim \delta) \ll 1/\epsilon$]. It is possible that the Bethe–Heitler (BH) situation takes place in this case. One can easily show that then $\delta \sim L\mu^2/2E$. Thus, in general, the diffusion regime occurs for the gluons with energy $\omega \gtrsim \omega_{cr}$, where

$$\omega_{cr} \sim \max\left(\frac{nC_3L^2}{4}, \frac{L\mu^2}{2}\right). \qquad (22)$$

Let us now discuss the energy loss. It can be written as

$$\Delta E = \int\limits_{\mu}^{\omega_{cr}} d\omega\,\omega\frac{dP}{d\omega} + \int\limits_{\omega_{cr}}^{\omega_{max}} d\omega\,\omega\frac{dP}{d\omega}. \qquad (23)$$

One can show that the first term in Eq. (23) is independent of energy and is on the order of ΔE_{BDMS} (1) for both LPM and Bethe–Heitler situations. At $E \longrightarrow \infty$, the energy loss is dominated by the second term in Eq. (23), which grows logarithmically with E. Then, using Eq. (18) to logarithmic accuracy, one can obtain in the high-energy limit

$$\Delta E = \frac{C_F\alpha_s}{4}\frac{L^2\mu^2}{\lambda_g}\log\frac{E}{\omega_{cr}}. \qquad (24)$$

Here, we have used that $\alpha_s^2\pi C_F C_T A(0)n/2\mu^2 = 1/\lambda_g$. Note that, since $L \gg 1/\mu$, from Eq. (22) it follows that always $\omega_{cr} \gg \mu$. The qualitative estimates (including the region $\omega \lesssim \omega_{cr}$) show that the appearance of ω_{cr} under the logarithm in Eq. (24), instead of μ in Eq. (2) for the RHIC conditions ($L \sim 4$ fm), can suppress the energy loss at $E \sim 10$ GeV by a factor of ~0.5. For the SPS conditions ($L \sim 2$ fm), the suppression is not strong (~0.7–0.8 at $E \sim 5$ GeV). The above estimates are obtained for the plasma temperature $T = 250$ MeV. Note

that the absence of ω_{cr} in the GLV prediction, Eq. (2), is connected with the neglect in [13] of the mass effects in evaluating the phase factor which controls the interference of the gluon emission from different points of the quark trajectory.

The above analysis is valid for the gluon emission from a fast quark as well. In this case, C_F in Eq. (24) should be replaced by $2C_A$ (the factor 2 comes from symmetry of the spectrum with respect to substitution $x \longleftrightarrow (1 - x)$).

REFERENCES

1. R. Baier, D. Schiff, and B. G. Zakharov, hep-ph/0002198 (2000) and references therein.
2. L. D. Landau and I. Ya. Pomeranchuk, Dokl. Akad. Nauk SSSR **92**, 535 (1953); **92**, 735 (1953).
3. A. B. Migdal, Phys. Rev. **103**, 1811 (1956).
4. R. Baier, Yu. L. Dokshitzer, A. H. Mueller, and D. Schiff, Nucl. Phys. B **531**, 403 (1998).
5. R. Baier, Yu. L. Dokshitzer, A. H. Mueller, *et al.*, Nucl. Phys. B **483**, 291 (1997); **484**, 265 (1997).
6. B. G. Zakharov, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 906 (1996) [JETP Lett. **63**, 952 (1996)].
7. B. G. Zakharov, Yad. Fiz. **61**, 924 (1998) [Phys. At. Nucl. **61**, 838 (1998)].
8. B. G. Zakharov, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 585 (1997) [JETP Lett. **65**, 615 (1997)].
9. B. G. Zakharov, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 181 (1999) [JETP Lett. **70**, 176 (1999)].
10. B. G. Zakharov, in *Proceedings of the at 33rd Rencontres de Moriond: QCD and High Energy Hadronic Interactions, Les Arcs, France, 1998*, Ed. by J. Tran Thanh Van, p. 533; hep-ph/9807396.
11. U. A. Wiedemann, Nucl. Phys. B **588**, 303 (2000).
12. M. Gyulassy and X.-N. Wang, Nucl. Phys. B **420**, 583 (1994).
13. M. Gyulassy, P. Levai, and I. Vitev, nucl-th/0005032 (2000).
14. N. N. Nikolaev and B. G. Zakharov, Zh. Éksp. Teor. Fiz. **105**, 1117 (1994) [JETP **78**, 598 (1994)].

On Compatibility of Tensor Products on Solutions to Commutativity and WDVV Equations[1]

A. Losev* and I. Polyubin**

* Institute for Theoretical and Experimental Physics, ul. Bol'shaya Cheremushkinskaya 25, Moscow, 117259 Russia

e-mail: lossev@vxitep.itep.ru

** Landau Institute for Theoretical Physics, Russian Academy of Sciences, Chernogolovka, 142432 Moscow region, Russia

e-mail: polyub@itp.ac.ru

Received December 28, 2000

We show that topological conformal theory contains solutions for WDVV equations for genus zero amplitudes and solutions to Commutativity equations from its quantum mechanics on the annulus. We explain behavior of these solutions under the tensor product of the theories. We make a conjecture that it is compatible with the construction of solutions to WDVV equations from a solution to the Commutativity equation and explicitly check it in the first nontrivial case. © *2001 MAIK "Nauka/Interperiodica"*.

PACS number: 11.25.Hf

The topological string theory (TST) is based on the 2D conformal topological theory that has nilpotent odd symmetry Q and Q-exact energy–momentum tensor: $T = Q(G)$. The n-point amplitude (traditionally denoted as F) in TST is given as an integral over the moduli space $M_{g,n}$ of complex structures on the genus g Riemann surface Σ_g with n-marked points. The integrand is the correlator of n Q-closed vertex operators Φ_{i_1}, …, Φ_{i_n} and $2n + 2g - 6$ superpartners of the energy–momentum tensor G:

$$
\begin{aligned}
F_{i_1 \ldots i_n} = \int_{M_{g,n}} &\left\langle \Phi_{i_1}(P_1) \ldots \Phi_{i_n}(P_n) \right. \\
&\times \prod_{a=1}^{3g+n-3} \left(\int_\Sigma \mu_a G_L dm^a \right) \left(\int_\Sigma \bar{\mu}_a G_R \overline{dm}^a \right) \Bigg\rangle,
\end{aligned}
\tag{1}
$$

where $G_L(G_R)$ is the superpartner of the energy–momentum tensor for the left (right) movers respectively and μ_a is the basis in the space of Beltrami differentials corresponding to coordinates m^a on the moduli space [1–5].

The notion of the amplitude in TST can be extended to the notion of Generalized Amplitude (GA) $GA_{i_1 \ldots i_n}$ [6, 7] taking values in cohomologies of the Deligne–Mumford compactification $\overline{M}_{g,n}$ of the moduli space of complex structures. Its value on the cycle $C_A \subset \overline{M}_{g,n}$ is given by the integral

$$
GA_{i_1 \ldots i_n}(C_A) = \int_{C_A} (\Omega)^{(\dim C_A)}_{i_1 \ldots i_n},
\tag{2}
$$

where $(\Omega)^{(\dim C_A)}_{i_1 \ldots i_n}$ is the dim C_A component of the closed form $\Omega_{i_1 \ldots i_n}$ given by the correlator

$$
\begin{aligned}
\Omega_{i_1 \ldots i_n} = &\left\langle \Phi_{i_1}(P_1) \ldots \Phi_{i_n}(P_n) \right. \\
&\times \exp\left(\int_\Sigma \sum_a \mu_a G_L dm^a \right) \exp\left(\int_\Sigma \sum_a \bar{\mu}_{\bar{a}} G_R d\overline{m}^{\bar{a}} \right) \Bigg\rangle.
\end{aligned}
\tag{3}
$$

S. Keel [8] found that the homology ring H_* of $\overline{M}_{0,k}$ is generated by cycles numbered by trees. These cycles correspond to the degeneration of a sphere into a collection of spheres joined by the double points with each other. Each sphere correspond to a vertex, and a point of junction of two spheres corresponds to an edge. Below, we will denote such vertices by parentheses containing labels of points and the labels of double points. For example, a cycle that corresponds to degeneration of the sphere with five marked points (such that

---

[1] This article was submitted by the authors in English.

points 1, 2, 3 belong to the first sphere and points 4, 5 belong to the second) is denoted as (1236)(745). Here, 6 and 7 are labels of the point of junction on the first and the second spheres, respectively.

Let us further assume that conformal dimensions of all fields in conformal topological theory are greater than or equal to zero, and the only fields that have zero dimensions are vertex operators Φ_i. Then, from the factorization properties of the correlators in conformal field theory, we conclude that the correlator on the degenerated surface is expressed in terms of correlators on components with insertions of the full basis of vertex operators at the place of the double point. For example, the correlator on a surface that is a junction of two spheres looks like

$$\Omega_{i_1 \dots i_n} = \Omega_{i_1 \dots i_{n_1} j} \Omega_{i_{n_1+1} \dots i_n k} \eta^{jk}, \qquad (4)$$

where η^{ij} is inverse to the bilinear pairing on vertex operators (given by a two-point correlator on the sphere) [1, 3, 9].

Keel described relations between cycles corresponding to trees. The evaluation of GA on these relations gives a set of constraints that are compactly represented as follows.

Introducing formal parameters T_i, we define the formal series

$$F(T) = \sum_{k=3}^{\infty} \frac{1}{k!} F_{i_1 \dots i_k} T_{i_1} \dots T_{i_k}. \qquad (5)$$

Keel's relations, together with the factorization conditions, lead to Associativity (WDVV) equations [2, 9, 3, 7] that were extensively studied in [10]:

$$\frac{\partial^3 F(T)}{\partial T_i \partial T_j \partial T_k} \eta^{kl} \frac{\partial^3 F(T)}{\partial T_l \partial T_p \partial T_q}$$

$$= \frac{\partial^3 F(T)}{\partial T_i \partial T_p \partial T_k} \eta^{kl} \frac{\partial^3 F(T)}{\partial T_l \partial T_j \partial T_q}. \qquad (6)$$

We say that the TST(I · II) is a product of theories I and II if its Lagrangian is a sum of Lagrangians. From this we get

$$Q^{\mathrm{I} \cdot \mathrm{II}} = Q^{\mathrm{I}} + Q^{\mathrm{II}};$$

$$G^{\mathrm{I} \cdot \mathrm{II}} = G^{\mathrm{I}} + G^{\mathrm{II}}; \quad \Phi_{\mathrm{I}}^{\mathrm{I} \cdot \mathrm{II}} = \Phi_i^{\mathrm{I}} \cdot \Phi_{\bar{i}}^{\mathrm{II}}. \qquad (7)$$

Thus, the differential form $\Omega_{I_1 \dots I_n}^{\mathrm{I} \cdot \mathrm{II}}$ in the product theory is a pointwise product (on the moduli space of complex structures) of differential forms Ω^{I} and Ω^{II}: $\Omega_{I_1 \dots I_n}^{\mathrm{I} \cdot \mathrm{II}} = \Omega_{i_1 \dots i_n}^{\mathrm{I}} \Omega_{\bar{i}_1 \dots \bar{i}_n}^{\mathrm{II}}$.

Therefore, the GA for the product of two theories is a product of GA in cohomologies. It can be computed

using the intersection product in homologies; in particular, the amplitude equals

$$F_{I_1 \dots I_n}^{\mathrm{I} \cdot \mathrm{II}} = (F^{\mathrm{I}} \cdot F^{\mathrm{II}})_{I_1 \dots I_n}$$

$$= (N^{-1})^{AB} GA_{i_1 \dots i_n}^{\mathrm{I}}(C_A) GA_{\bar{i}_1 \dots \bar{i}_n}^{\mathrm{II}}(C_B), \qquad (8)$$

where C_A is the basis in the space of homologies, $N_{AB} = C_A \cap C_B$. Due to factorization property (4), the GA can be expressed in terms of F; thus, Eq. (8) can be considered as a formal product of solutions to WDVV equations (even if these solutions are not coming from any TST) [11].

It is believed that the above scheme of reasoning (with some modifications) can be applied to the wider class of theories, for example, to theories that can contain gravitational descendants (that lead to Morita–Mumford classes on the Deligne–Mumford compactification of moduli space). In particular, action in such theory can be deformed by adding to it 2-observables, corresponding to gravitational descendants. As was shown in [12–14, 11], the generalized amplitudes in these theories also can be considered as maps to cohomologies of $\overline{M}_{0,n}$ and, thus, lead to solutions of WDVV equations. So the product operation for GA can be defined in these theories as above.

Each 2D topological theory contains a supersymmetric quantum mechanics whose Hilbert space $\mathcal{H}$ is a space of wave functions associated to the circle (that are invariant under the rotations of this circle), with supersymmetries and Hamiltonian [15, 16]

$$Q_{\pm} = Q_{L,0} \pm Q_{R,0}, \quad G_{\pm} = G_{L,0} \pm G_{R,0},$$

$$H = T_{L,0} + T_{R,0}$$

that obey (anti)commutation relations

$$\{Q_+, G_+\} = \{Q_-, G_-\} = H. \qquad (9)$$

This quantum mechanics also contains Q_+-closed operators $\phi_i = \int_{x \in S_1} \Phi(x) dx$, where $0 \le x \le 2\pi$ is a standard coordinate on S_1.

We assume that the kernel W of the Hamiltonian consists of Q_+ and G_- closed states, and the rest of its eigenvalues have a strictly positive real part.

One can show that the regularity of the form $\Omega_{i_1 \dots i_n}$ (3) at points of degeneration of Riemann surface leads to

$$[\phi_i, \phi_j] = [\phi_i, [G_-, \phi_j]] = 0. \qquad (10)$$

Given N operators $\phi_{i_1}, \dots, \phi_{i_N}$, we define the form $\omega_{i_1, \dots, i_N}$ on $(C^*)^N / C^*$ taking values in linear operators in W as follows.

Let us parameterize points on $C*$ by two standard real coordinates $l > 0$ and φ (taking the value in the interval $[0, 2\pi]$) : $z = \exp(l + i\varphi)$. Consider the space of N points on $(C*)^N$—it is parameterized by coordinates l_i and φ_i, $i = 1, \ldots, N$.

Let us delete all l diagonals ($l_i = l_j$ for some pair (i, j)) from the space $(C*)^N$. The rest of the space is divided into regions labeled by permutations σ of the first N integers, such that

$$l_{\sigma(k)} > l_{\sigma(m)} \, iff \, k > m. \tag{11}$$

The form $\omega_{i_1 \ldots i_N}$ is defined separately in each region:

$$\omega_{i_1 \ldots i_N} = \Pi_W \Phi_{i_{\sigma(1)}} \exp(\hat{H}_{\sigma(1)\sigma(2)}) \Phi_{i_{\sigma(2)}}$$
$$\times \exp(\hat{H}_{\sigma(2)\sigma(3)}) \ldots \Phi_{i_{\sigma(N)}} \Pi_W, \tag{12}$$

where

$$\hat{H}_{km} = -H(l_k - l_m)$$
$$+ G_+(dl_k - dl_m) + G_-(d\varphi_k - d\varphi_m) \tag{13}$$

and Π_W is the projector on W.

One can easily check that in each region the form is closed: $Q_+\Pi_W = 0$ and

$$d(\hat{H}_{km}) = [Q_+, \hat{H}_{km}]. \tag{14}$$

Now, we would like to get the closed form on the union of regions: we need to take care about discontinuities. Suppose that at the boundary of the regions $l_k = l_m$. Then, discontinuities in the terms containing $d(l_k - l_m)$ are irrelevant—these terms are closed anyway. So, we have to take care about terms containing no differentials in the $l_k - l_m$ direction and about terms containing $d(\varphi_k - \varphi_m)$, but these discontinuities are absent due to Eq. (10).

Now we come to the issue of the "sausage" compactification: the proper compactification of $(C*)^N/C*$ is such that $\omega_{i_1 \ldots i_N}$ is defined and closed on it.

The obvious "infinities" are directions when $(l_{\sigma(p+1)} - l_{\sigma(p)}) \longrightarrow +\infty$, where $\exp(-(l_{\sigma(p+1)} - l_{\sigma(p)})H) \longrightarrow \Pi_W$.

The component of the differential form $\omega_{i_1 \ldots i_N}$ that does contain neither $d\varphi_{\sigma(p+1)} - d\varphi_{\sigma(p)}$ nor $dl_{\sigma(p+1)} - dl_{\sigma(p)}$ tends to its factorized value, while components that contain any of them rapidly tend to zero, because $G_+\Pi_W = G_-\Pi_W = 0$.

The proper compactification of $(C*)^{\otimes n}/C*$ was introduced and studied in [17, 18]. In particular, it was shown that the set of forms $\omega_{i_1 \ldots i_n}$ are maps to cohomologies of the new moduli spaces $\bar{L}_n$ taking values in linear operators in W that satisfy the factorization property. It was shown that (similar to $\bar{M}_{0,n}$ case) all

cycles in $\bar{L}_n$ can be represented as a sequence of spheres containing marked points and special points 0 and ∞, such that the ∞ of the kth sphere is attached to the 0 of the $k + 1$ sphere. Below, we will denote a sphere with marked points by parentheses containing labels of these points, and the sequence of spheres, by a sequence of such parentheses.

Relations between cycles in the homology group of $\bar{L}_n$ spaces were studied in [18], where it was shown that all these relations are generated as pullbacks of the Commutativity relation

$$(1)(2) = (2)(1) \tag{15}$$

under the forgetful map $\bar{L}_n \longrightarrow \bar{L}_2$ that forgets the positions of $(n - 2)$ points.

From the factorization property of ω discussed above, we see that, if the sequence S_1S_2 is a junction of two sequences S_1 and S_2, then

$$\int_{S_1 S_2} \omega_{i_1 \ldots i_{m+n}} = \int_{S_1} \omega_{i_1 \ldots i_m} \int_{S_2} \omega_{i_{m+1} \ldots i_{m+n}}. \tag{16}$$

Evaluation of ω on cycles that are relations in cohomology gives equations for

$$\tau_{i_1 \ldots i_m} = \int_{\bar{L}_m} \omega_{i_1 \ldots i_m}.$$

If we consider the generating function

$$\tau(t) = \sum_m \frac{1}{m!} \tau_{i_1 \ldots i_m} t^{i_1} \ldots t^{i_m}, \tag{17}$$

these equations would be equivalent to the Commutativity equations (known also as a t part of $t - t*$ equations [19]:

$$\frac{\partial \tau}{\partial t^i} \frac{\partial \tau}{\partial t^j} = \frac{\partial \tau}{\partial t^j} \frac{\partial \tau}{\partial t^i}. \tag{18}$$

Now consider the product of two theories:

$$\mathscr{H}^{I \cdot II} = \mathscr{H}^I \otimes \mathscr{H}^{II}, \quad Q^{I \cdot II} = Q^I + Q^{II},$$
$$G^{I \cdot II} = G^I + G^{II}, \tag{19}$$

$$\phi_I^{I \cdot II} = \phi_i^I \otimes \phi_i^{II}. \tag{20}$$

The form $\omega^{I \cdot II}$ on the space $\bar{L}_N$ is

$$\omega^{I \cdot II}_{I_1, \ldots, I_N} = \omega^I_{i_1, \ldots, i_N} \otimes \omega^{II}_{\bar{i}_1, \ldots \bar{i}_N}. \tag{21}$$

Thus,

$$\int_{\bar{L}_N} \omega^{I \cdot II}_{I_1, \ldots, I_N} = (n^{-1})^{AB} \int_{C_A} \omega^I_{i_1, \ldots, i_N} \otimes \int_{C_B} \omega^{II}_{\bar{i}_1, \ldots \bar{i}_N}. \tag{22}$$

Here, classes of C_A form a basis in the space of homologies of $\bar{L}_N$, and n is the intersection matrix in homologies of $\bar{L}_N$.

Due to the factorization property, $\int_{C_A} \omega_{i_1 \ldots i_N}$ can be expressed in terms of products of $\int_{L_m} \omega_{i_1 \ldots i_m}$; thus, $\tau^{\mathrm{I} \cdot \mathrm{II}}$ can be expressed in terms of τ^{I} and τ^{II}, $\tau^{\mathrm{I} \cdot \mathrm{II}} = \tau^{\mathrm{I}} \cdot \tau^{\mathrm{II}}$. As in the case of product of solutions to WDVV equations, Eq. (22) can be considered as a formal product operation on solutons to Commutativity equations (even if they are not coming from any TST). In [20], it was proposed that, starting with a solution to the Commutativity equation, one can construct a solution to the Associativity equation as follows. We pick up a vector $h \in W$ and make a change of coordinates from t to θ^a:

$$\theta^a(t) = \tau_b^a(t) h^b. \tag{23}$$

If matrix τ is symmetric with respect to some scalar product $\langle, \rangle$ and the change of coordinates is invertible, then

$$\tau_{ab} = \partial^2 F / \partial \theta^a \partial \theta^b \tag{24}$$

and F solves the Associativity equation with the metric $\eta_{ab} = \delta_{ab}$.

We will denote the construction above as the CTA map from the pair τ, h (such that τ solves the Commutativity equation) to $F(\theta)$ (such that F solves the Associativity equations): $F = \mathrm{CTA}(\tau, h)$.

Heuristically, vector h should be considered as a state that is generated by the functional integral over a disk in the theory with two-observable (corresponding to the first gravitational descendant) added to the action (in particular, it is well known [19] that adding to the action the first gravitational descendant of 1 corresponds to the rescaling of h).

The above heuristic considerations lead to the conjecture that can be rigorously checked and is really checked below in the first nontrivial case (corresponding to $\bar{M}_{0,5}$ and $\bar{L}_3$), that the CTA map is compatible with the product operation

$$\begin{aligned} &\mathrm{CTA}(\tau^{\mathrm{I}} \cdot \tau^{\mathrm{II}}, h^{\mathrm{I}} \otimes h^{\mathrm{II}}) \\ &= \mathrm{CTA}(\tau^{\mathrm{I}}, h^{\mathrm{I}}) \cdot \mathrm{CTA}(\tau^{\mathrm{II}}, h^{\mathrm{II}}). \end{aligned} \tag{25}$$

It would extremely desirable to have a complete proof of the conjecture (25), but at the moment the direct check looks very complicated.

In preparation to the check of this conjecture, we will make the linear parameter redefinition ($\theta \longrightarrow T$)

$$\theta_a = T^i \frac{\partial \tau_{ab}}{\partial t^i} \bigg|_{t=0} h^b.$$

In these coordinates, the metric equals $\eta_{ij} = \langle h \tau_i \tau_j h \rangle$.

Let us get an explicit expression for the rhs of Eq. (25). The five-point amplitude looks like

$$\begin{aligned} F_{ijlkm} &= \langle h \tau_i [\tau_{jkl}, \tau_m] h \rangle + \langle h [\tau_{im}, \tau_j] \tau_{kl} h \rangle \\ &+ \langle h [\tau_{im}, \tau_k] \tau_{jl} h \rangle + \langle h [\tau_{im}, \tau_l] \tau_{jk} h \rangle. \end{aligned} \tag{26}$$

Let us choose the basis in the space of cycles such that the intersection matrix is nondegenerate (it has rank 5):

$$\begin{aligned} &(ik)(jlm) + (jl)(ikm) \\ &+ (jm)(ikl); (ij)(klm); (jk)(ilm); (jl)(ikm); (jm)(ikl). \end{aligned} \tag{27}$$

In this basis, the intersection matrix is diagonal: $\mathrm{diag}(1; -1; -1; -1; -1)$.

The five-point amplitude for the tensor product of two theories in terms of amplitudes of these theories can be symbolically rewritten as

$$\begin{aligned} F_{(5)}^{1 \otimes 2} &= F_{(5)}(\bar{F}_{(3)} \bar{F}_{(3)} \bar{F}_{(3)}) + (F_{(3)} F_{(3)} F_{(3)}) \bar{F}_{(5)} \\ &+ (F_{(4)} F_{(3)})_a (\bar{F}_{(3)} \bar{F}_{(4)})_b (N^{-1})^{ab}, \end{aligned} \tag{28}$$

where $F_{(n)}$ denotes $F_{i_1 \ldots i_n}$, the first two terms correspond to intersection of cycles of real dimension 4 and 0, and the last term corresponds to intersection of cycles of real dimension 2; $F_{(i)}$; $\bar{F}_{(i)}$ correspond to the first and second theory correspondingly. Using Eq. (28), we get

$$\begin{aligned} F_{IJKLM}^{1 \otimes 2} &= F_{ijklm} \langle \overline{h \tau_i \tau_j \tau_k \tau_l \tau_m h} \rangle + \langle h \tau_i \tau_j \tau_k \tau_l \tau_m h \rangle \bar{F}_{ijklm} \\ &- \langle h \tau_i \tau_j [\tau_{lm}, \tau_k] h \rangle \langle \overline{h \tau_i \tau_j [\tau_{lm}, \tau_k] h} \rangle \\ &- \langle h \tau_j \tau_k [\tau_{lm}, \tau_i] h \rangle \langle \overline{h \tau_j \tau_k [\tau_{lm}, \tau_i] h} \rangle \\ &- \langle h \tau_j \tau_l [\tau_{ik}, \tau_m] h \rangle \langle \overline{h \tau_j \tau_l [\tau_{ik}, \tau_m] h} \rangle \\ &- \langle h \tau_j \tau_m [\tau_{ik}, \tau_l] h \rangle \langle \overline{h \tau_j \tau_m [\tau_{ik}, \tau_l] h} \rangle \\ &+ (\langle h \tau_i \tau_k [\tau_{lm}, \tau_j] h \rangle + \langle h \tau_j \tau_l [\tau_{ik}, \tau_m] h \rangle \\ &+ \langle h \tau_j \tau_m [\tau_{ik}, \tau_l] h \rangle)(\langle \overline{h \tau_i \tau_k [\tau_{lm}, \tau_j] h} \rangle \\ &+ \langle \overline{h \tau_j \tau_l [\tau_{ik}, \tau_m] h} \rangle + \langle \overline{h \tau_j \tau_m [\tau_{ik}, \tau_l] h} \rangle). \end{aligned} \tag{29}$$

For $\bar{L}_3$ moduli space, the intersection matrix of 2-cycles has rank 4. In the basis of cycles

$$\begin{aligned} &(jl)(k) + (j)(kl) \\ &+ (jk)(l); (jk)(l); (jl)(k); (kl)(j), \end{aligned} \tag{30}$$

the intersection matrix is diagonal: $\text{diag}(1; -1; -1; -1)$. The lhs of Eq. (25) can be written as

$$F_{ijklm}^{1\otimes 2} = \langle h^{1\otimes 2}\tau_i^{1\otimes 2}[\tau_{jkl}^{1\otimes 2}, \tau_m^{1\otimes 2}]h^{1\otimes 2}\rangle$$
$$+ \langle h^{1\otimes 2}[\tau_{im}^{1\otimes 2}, \tau_j^{1\otimes 2}]\tau_{kl}^{1\otimes 2}h^{1\otimes 2}\rangle$$
$$+ \langle h^{1\otimes 2}[\tau_{im}^{1\otimes 2}, \tau_k^{1\otimes 2}]\tau_{jl}^{1\otimes 2}h^{1\otimes 2}\rangle \qquad (31)$$
$$+ \langle h^{1\otimes 2}[\tau_{im}^{1\otimes 2}, \tau_l^{1\otimes 2}]\tau_{jk}^{1\otimes 2}h^{1\otimes 2}\rangle,$$

where

$$h^{1\otimes 2} = h\otimes\bar{h}, \quad \tau_i^{1\otimes 2} = \tau_i\otimes\bar{\tau}_i,$$

$$\tau_{jk}^{1\otimes 2} = \tau_{jk}\otimes\overline{\tau_j\tau_k} + \tau_j\tau_k\otimes\overline{\tau_{jk}},$$

$$\tau_{jkl}^{1\otimes 2} = \tau_{jkl}\otimes\overline{\tau_j\tau_k\tau_l} + \tau_j\tau_k\tau_l\otimes\overline{\tau_{jkl}}$$
$$+ \tau_{kl}\tau_j\otimes\overline{\tau_{kl}\tau_j} + \tau_{kl}\tau_j\otimes\overline{\tau_l\tau_{jk}} \qquad (32)$$
$$+ \tau_{kl}\tau_j\otimes\overline{\tau_k\tau_{jl}} + \tau_k\tau_{jl}\otimes\overline{\tau_{kl}\tau_j}$$
$$+ \tau_k\tau_{jl}\otimes\overline{\tau_l\tau_{jk}} + \tau_l\tau_{jk}\otimes\overline{\tau_{kl}\tau_j}$$
$$+ \tau_l\tau_{jk}\otimes\overline{\tau_k\tau_{jl}} - \tau_j\tau_{kl}\otimes\overline{\tau_j\tau_{kl}}.$$

Substituting Eq. (32) into Eq. (31), one gets

$$F_{ijklm}^{1\otimes 2} = \langle h\tau_j[\tau_{ilm}, \tau_k]h\rangle\langle\overline{h\tau_i\tau_j\tau_k\tau_l\tau_m h}\rangle$$

$$+ \langle h\tau_i\tau_j\tau_k\tau_l\tau_m h\rangle\langle\overline{h\tau_j[\tau_{ilm}, \tau_k]h}\rangle$$

$$+ \langle h\tau_i\tau_j\tau_k\tau_{lm}h\rangle\langle\overline{h\tau_i\tau_j\tau_k\tau_{lm}h}\rangle$$

$$- \langle h\tau_i\tau_j\tau_{lm}\tau_k h\rangle\langle\overline{h\tau_i\tau_j\tau_{lm}\tau_k h}\rangle$$

$$+ \langle h\tau_j\tau_k\tau_m\tau_{il}h\rangle\langle\overline{h\tau_j\tau_k\tau_m\tau_{il}h}\rangle$$

$$- \langle h\tau_j\tau_m\tau_{il}\tau_k h\rangle\langle\overline{h\tau_j\tau_m\tau_{il}\tau_k h}\rangle$$

$$+ \langle h\tau_j\tau_k\tau_l\tau_{im}h\rangle\langle\overline{h\tau_j\tau_k\tau_l\tau_{im}h}\rangle$$

$$- \langle h\tau_j\tau_l\tau_{im}\tau_k h\rangle\langle\overline{h\tau_j\tau_l\tau_{im}\tau_k h}\rangle$$

$$+ (\langle h\tau_j\tau_{lm}\tau_i\tau_k h\rangle + \langle h\tau_j\tau_l\tau_{im}\tau_k h\rangle + \langle h\tau_j\tau_m\tau_{il}\tau_k h\rangle)$$

$$\times (\overline{\langle h\tau_j\tau_{lm}\tau_i\tau_k h\rangle + \langle h\tau_j\tau_l\tau_{im}\tau_k h\rangle + \langle h\tau_j\tau_m\tau_{il}\tau_k h\rangle})$$

$$- (\langle h\tau_j\tau_k\tau_{lm}\tau_i h\rangle + \langle h\tau_j\tau_k\tau_l\tau_{im}h\rangle + \langle h\tau_j\tau_k\tau_m\tau_{il}h\rangle)$$

$$\times (\overline{\langle h\tau_j\tau_k\tau_{lm}\tau_i h\rangle + \langle h\tau_j\tau_k\tau_l\tau_{im}h\rangle + \langle h\tau_j\tau_k\tau_m\tau_{il}h\rangle})$$

$$+ \langle h[\tau_{jk}, \tau_i]\tau_{lm}h\rangle\langle\overline{h\tau_i\tau_j\tau_k\tau_l\tau_m h}\rangle$$

$$+ \langle h\tau_i\tau_j\tau_k\tau_l\tau_m h\rangle\langle\overline{h[\tau_{jk}, \tau_i]\tau_{lm}h}\rangle \qquad (33)$$

$$+ \langle h\tau_i\tau_j\tau_k\tau_{lm}h\rangle\langle\overline{h[\tau_{jk}, \tau_i]\tau_l\tau_m h}\rangle$$

$$+ \langle h[\tau_{jk}, \tau_i]\tau_l\tau_m h\rangle\langle\overline{h\tau_i\tau_j\tau_k\tau_{lm}h}\rangle$$

$$+ \langle h[\tau_{jk}, \tau_l]\tau_{im}h\rangle\langle\overline{h\tau_i\tau_j\tau_k\tau_l\tau_m h}\rangle$$

$$+ \langle h\tau_i\tau_j\tau_k\tau_l\tau_m h\rangle\langle\overline{h[\tau_{jk}, \tau_l]\tau_{im}h}\rangle$$

$$+ \langle h\tau_j\tau_k\tau_l\tau_{im}h\rangle\langle\overline{h[\tau_{jk}, \tau_l]\tau_i\tau_m h}\rangle$$

$$+ \langle h[\tau_{jk}, \tau_l]\tau_i\tau_m h\rangle\langle\overline{h\tau_j\tau_k\tau_l\tau_{im}h}\rangle$$

$$+ \langle h[\tau_{jk}, \tau_m]\tau_{il}h\rangle\langle\overline{h\tau_i\tau_j\tau_k\tau_l\tau_m h}\rangle$$

$$+ \langle h\tau_i\tau_j\tau_k\tau_l\tau_m h\rangle\langle\overline{h[\tau_{jk}, \tau_m]\tau_{il}h}\rangle$$

$$+ \langle h\tau_j\tau_k\tau_m\tau_{il}h\rangle\langle\overline{h[\tau_{jk}, \tau_m]\tau_i\tau_l h}\rangle$$

$$+ \langle h[\tau_{jk}, \tau_m]\tau_i\tau_l h\rangle\langle\overline{h\tau_j\tau_k\tau_m\tau_{il}h}\rangle.$$

Straightforward but tedious calculations show that Eq. (33) is Eq. (29), so (25) is fulfilled.

In this paper, we motivated and checked the compatibility between product operations on solutions to Commutativity and Associativity equations in the first nontrivial case. It is not excluded that in more complicated cases these two product would give different results. If this happens, it would give a nontrivial indicator of whether solution to Commutativity equations came from TST or not. At the same time, if conjecture turns out to be a theorem, it would drastically simplify the calculation of the product on solutions to Associativity equations (because spaces $\bar{L}_n$ are much more tractable than $\bar{M}_{0,n}$).

REFERENCES

1. E. Witten, Nucl. Phys. B **340**, 281 (1990).
2. R. Dijkgraaf and E. Witten, Nucl. Phys. B **342**, 486 (1990).
3. R. Dijkgraaf, E. Verlinde, and H. Verlinde, in *Proceedings of the Trieste Spring School, 1990*, Ed. by M. Green *et al.* (World Scientific, Singapore, 1991).
4. A. Losev, Theor. Math. Phys. **95**, 595 (1993).
5. T. Eguchi, H. Kanno, Y. Yamada, *et al.*, Phys. Lett. B **305**, 235 (1993).
6. E. Witten, hep-th/9207094.
7. M. Kontsevich and Yu. Manin, Commun. Math. Phys. **164** (3), 525 (1994).
8. S. Keel, Trans. Am. Math. Soc. **330**, 545 (1992).
9. E. Verlinde and H. Verlinde, Nucl. Phys. B **348**, 457 (1991).
10. B. Dubrovin, Lect. Notes Math. **1620**, 120 (1996).

11. M. Kontsevich and Yu. Manin, with appendix by Kaufmann, q-alg/9502009.

12. I. M. Krichever, in *30 Years of the Landau Institute: Selected Papers*, Ed. by I. M. Khalatnikov and V. P. Mineev (World Scientific, Singapore, 1996); hep-th/9205110.

13. A. Losev, in *Proceedings of the 3rd Baltic Rim Student Seminar on Integrable Models and Strings, Helsinki, 1993*, Ed. by A. Alekseev (Springer-Verlag, Berlin, 1995), p. 172.

14. A. Losev and I. Polyubin, Int. J. Mod. Phys. A **10**, 4161 (1995).

15. S. Cecotti, L. Girardello, and A. Pasquinicci, Int. J. Mod. Phys. A **6**, 2427 (1991).

16. S. Cecotti, Nucl. Phys. B **335**, 755 (1991).

17. A. Losev, Preprint ITEP-TH-84/98, LPTHE-61/98.

18. A. Losev and Yu. Manin, Mich. Math. J. **48**, 443 (2000); alg-geom/0001003.

19. S. Cecotti and C. Vafa, Nucl. Phys. B **367**, 359 (1991).

20. A. Losev, in *Proceedings of 38th Taniguchi Symposium "Topological Field Theory, Primitive Form and Related Topics," 1996* (Birkhauser, Boston, 1998); hep-th/9801179.

Nonlinear Bloch Waves[1]

V. N. Serkin, M. Matsumoto*, and T. L. Belyaeva**

Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 39, Moscow, 117942 Russia

Benemerita Universidad Autonoma de Puebla, 72001 Puebla, Mexico

e-mail: vserkin@hotmail.com

* Osaka University, Osaka 565-0871, Japan

** Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

Received November 2, 2000

The nonlinear Bloch theorem for the temporal and spatial Schrödinger solitons in dispersive and nonlinear periodic structures is proved. It is shown that bright and dark solitary nonlinear Bloch waves exist only under certain conditions and that the parameter functions describing dispersion and nonlinearity periodic inhomogeneities cannot be chosen independently. © *2001 MAIK "Nauka/Interperiodica".*

PACS numbers: 42.65.Tg; 42.81.Dp; 05.45.Yv

A fundamental theorem concerning electrons in a crystal was proved by F. Bloch in 1928 [1]. It was shown that the wave functions of electrons in a periodic crystal lattice have the Bloch form

$$\Psi_k(x) = U_k(x)\exp(ikx), \qquad (1)$$

where the function $U_k(x) = U_k(x + L)$ possesses the same periodicity as the lattice, and the wave number k is related to the de Broglie wavelength of the electron. The Bloch function in one-dimensional form (1) corresponds to a free electron wave function, $\exp(ikx)$, modulated by a function $U_k(x)$, which has the periodicity of the crystal lattice. The Bloch theorem is one of the basic concepts in solid state physics (see, for example, [2–6]).

In this letter, we show that there exists a nonlinear Bloch theorem for temporal and spatial Schrödinger solitons propagating through an inhomogeneous nonlinear and dispersive structures characterized by translational symmetry.

Let us consider the nonlinear Schrödinger equation (NLSE) model with periodically varying nonlinearity and dispersion:

$$i\frac{\partial \Psi^{\pm}}{\partial Z} \pm \frac{1}{2}D(Z)\frac{\partial^2 \Psi^{\pm}}{\partial X^2} + R(Z)|\Psi^{\pm}|^2\Psi^{\pm} = 0. \qquad (2)$$

Equation (2) is written here in the standard soliton units, as they are commonly known. It is assumed that periodic perturbations to the dispersion $D(Z) = D(Z + nL)$ and nonlinearity $R(Z) = R(Z + nL)$ are not limited to

the regime where they are smooth and small. Here, we also assume that the periods of these functions are arbitrary values. Due to a spatial–temporal analogy, both temporal and spatial solitons are described by Eq. (2).

Nonlinear Bloch theorem. The transformation law for the bright $\Psi^+(Z, X)$ and dark $\Psi^-(Z, X)$ NLSE solitons propagating through an inhomogeneous medium characterized by translational symmetry $D(Z) = D(Z + nL)$ and/or $R(Z) = R(Z + nL)$ is defined by the self-reproducing stable configuration

$$\Psi^{\pm}(Z) = \sqrt{P(Z)}\begin{Bmatrix} \sqrt{C}\eta\,\mathrm{sech}(\eta P(Z)X) \\ \sqrt{C}\eta\,\tanh(\eta P(Z)X) \end{Bmatrix}$$
$$\times \exp\left[\pm i\frac{P(Z)}{2}X^2 + i\int_0^Z K^{\pm}(\zeta)d\zeta\right], \qquad (3)$$

where the real function $P(Z)$ possesses the same periodicity as the medium: $P(Z) = P(Z + nL)$. Nonlinear Bloch waves exist only under certain conditions given by the following relations:

$$P(Z) = P_0\left[1 + P_0\int_0^Z \Phi(Z')dZ'\right]^{-1};$$
$$R(Z) = P(Z)\Phi(Z)/C, \qquad (4)$$

where the dispersion parameter function is assumed to be given in the form of a periodic function $D(Z + nL) = \Phi(Z + nL)$.

---

[1] This article was submitted by the authors in English.

In another case, when the nonlinearity is assumed to be a given periodic function $R(Z) = R(Z + nL)$, the nonlinear Bloch wave exists only under conditions

$$P(Z) = P_0 \exp\left[-C\int_0^Z R(Z')dZ'\right];$$

$$D(Z) = CR(Z)/P(Z). \tag{5}$$

Parameter $C = D_0 P_0/R_0$ in Eqs. (4) and (5) is determined by the initial conditions at $Z = 0$.

Proof. Substitution of the nonlinear Bloch's function

$$\Psi^\pm(Z) = \sqrt{P(Z)}Q^\pm(S)$$

$$\times \exp\left[\pm i\frac{P(Z)}{2}X^2 + i\int_0^Z K^\pm(\zeta)d\zeta\right] \tag{6}$$

in Eq. (2) leads to

$$\pm\frac{1}{2}\frac{\partial^2 Q^\pm}{\partial S^2} + \frac{R}{DP}(Q^\pm)^3$$

$$-Q^\pm\frac{K^\pm}{DP^2} \mp \frac{S^2 Q^\pm}{2DP^4}\left(DP^2 + \frac{\partial P}{\partial Z}\right) = 0, \tag{7}$$

$$\left(DP^2 + \frac{\partial P}{\partial Z}\right)\left(\frac{1}{2}Q^\pm + S\frac{\partial Q^\pm}{\partial S}\right) = 0. \tag{8}$$

One can transform the general system of Eqs. (7), (8) to an exactly integrable form. Let us consider the complete nonlinear regime, when Eq. (7) represents the "classical," exactly integrable NLSE model and when two Eqs. (8) have a nonsingular solution for the Q function. The necessary transformation can be expressed by the following system of equations:

$$P^{-1}(Z) = C + \int D(\zeta)d\zeta;$$

$$D(Z)P(Z) = R(Z)C. \tag{9}$$

The eigenvalues $E^\pm = K^\pm/DP^2$ in Eq. (7) must satisfy $E^+ = 0.5\eta^2$ and $E^- = \eta^2$ for the bright and dark nonsingular soliton solutions of Eqs. (7), (8):

$$Q(S)^\pm = \begin{cases} \eta\,\mathrm{sech}\,(\eta P(Z)X) \\ \eta\,\mathrm{tanh}\,(\eta P(Z)X) \end{cases}. \tag{10}$$

So, it is not surprising to find from Eqs. (9) that the variation of soliton shape is also periodic, $P(Z) = P(Z + nL)$, if the following relations are satisfied:

$$\frac{R(Z + nL)}{R(Z)} = \frac{D(Z + nL)}{D(Z)}. \tag{11}$$

Notice that the necessary condition for the existence of a Bloch theorem in the NLSE model with periodic gain or loss $\Gamma(Z) = \Gamma(Z + nL)$ is

$$D(Z)P(Z) = R(Z)C\exp\left(2\int_0^Z \Gamma(\xi)d\xi\right). \tag{12}$$

It should also be noted that exact integrability of the model (2) under conditions (9) means that there exists a transformation law (nonlinear Bloch theorem) for the high-order soliton solutions of Eq. (2) as well.

Two features of the exact self-reproducing stable solutions (3)–(5) and (9) are noteworthy.

(1) Solutions (3)–(5) and (9) can be considered as nonlinear Bloch waves with a periodic scattering potential, which is reproduced by a solitary wave itself from cell to cell of a periodic structure. Unlike the homogeneous medium solution [$D(Z) \equiv 1$ and $R(Z) \equiv 1$], the soliton amplitude, pulse width and chirp (for the case of temporal solitons), or the radius of curvature of the wavefront (for the case of spatial solitons) has periodicity of the medium symmetry.

(2) The dependence of soliton pulse width and phase profile on the propagation distance is defined by the same periodic function $P(Z)$. This remarkable result opens up the possibility to construct different nonlinear Bloch functions by using the algorithm proposed.

The fundamental set of nonlinear Bloch waves can be represented by Jacobi elliptic functions:

$$P(Z) = A + Bdn^2(Z; s);$$

$$D(Z) = \frac{s^2 Bsn(2Z; s)[1 - s^2 sn^4(Z; s)]}{[A + Bdn^2(Z; s)]^2}, \tag{13}$$

where $dn^2 = 1 - s^2 sn^2(Z; s)$, and $sn(Z; s)$ is the elliptical sine function with the module parameter $m = s^2$. Periodic function $P(Z)$, represented by Eq. (13), has a period equal to $2K(s)$, where $K(s)$ is a complete elliptic integral of the first kind [6]. Periodic function $P(Z)$ (13) transforms to a sinelike wave for parameter $s \ll 1$:

$$P(Z) = A + B(1 - s^2\sin^2 Z);$$

$$D(Z) = s^2 B\sin(2Z) \tag{14}$$

[with asymptotic period given by $2K(s) = \pi(1 - s^2/4)$ for $s \ll 1$]; and it transforms to a periodic train of sechlike solitons for $s \longrightarrow 1$:

$$P(Z) = A + B\,\mathrm{sech}^2(Z);$$

$$D(Z) = \frac{2s^2 B\,\mathrm{sech}^2(Z)\tanh(Z)}{[A + B\,\mathrm{sech}^2(Z)]^2},$$

with period given by asymptotic value $2K(s) = \ln(16/(1 - s^2))$ for $s \longrightarrow 1$.

In the case when a periodic structure is closed in the form of a loop with a total length equal to $Z = nL$, such
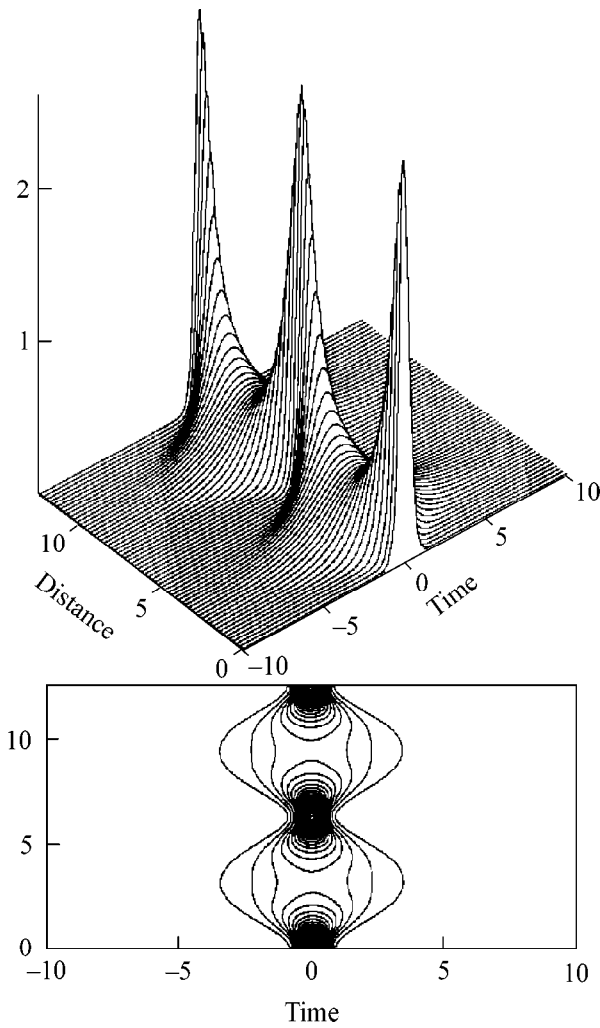
Fig. 1. Evolution of the nonlinear solitary Bloch wave [Eqs. (3), (15), and (16)] as a function of the propagation distance. Initial conditions: $k = 1$ and $\beta = 0.5$.
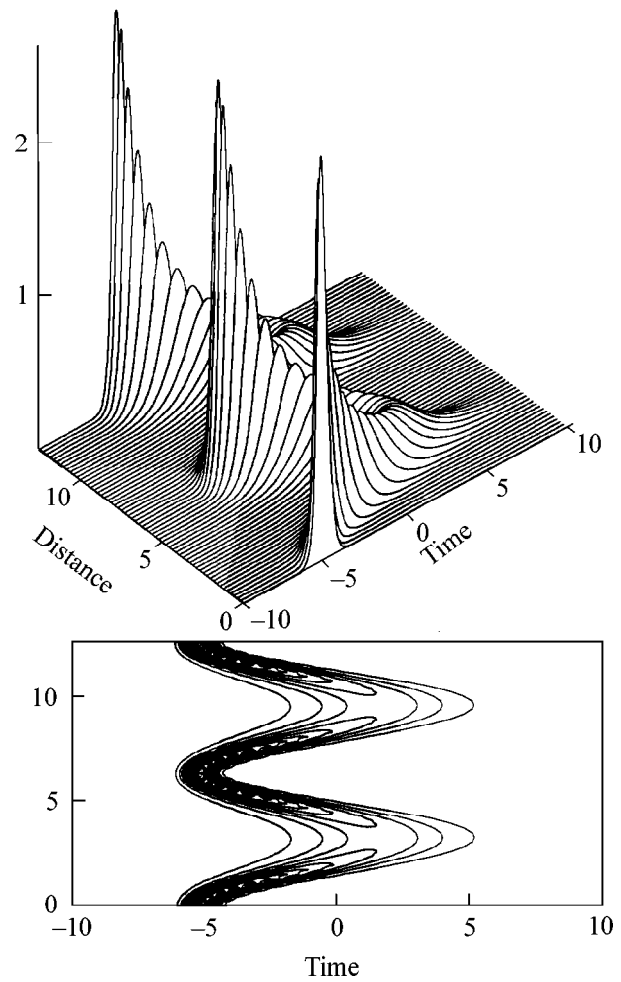

Fig. 2. Evolution of the nonlinear solitary Bloch wave [Eqs. (3), (15), and (16)] with nontrivial initial conditions as a function of the propagation distance: soliton "snake" effect. Initial conditions: initial group velocity $V = 10$; $k = 1$; and $\beta = 0.5$.

a structure can be considered as an example of the soliton memory [7]. One of the simplest periodic solutions of Eq. (2) in this case is given by a periodic function $P(Z)$:

$$P(Z) = \frac{\sqrt{1 - \beta^2}}{1 - \beta \cos kZ}, \tag{15}$$

which, for $\beta \longrightarrow 1$, becomes a periodic grating of delta functions and, for $\beta \longrightarrow 0$, a constant $P(Z) \longrightarrow 1$. Then, the nonlinear Bloch theorem (3) leads to the requirement that dispersion and nonlinearity of the soliton memory loop must be given by periodic functions:

$$D(Z) = \frac{\beta k \sin kZ}{\sqrt{1 - \beta^2}}; \quad R(Z) = \frac{\beta k \sin kZ}{1 - \beta \cos kZ}. \tag{16}$$

It seems very attractive to use the nonlinear Bloch concept to design novel types of soliton lasers. The most noteworthy feature of the quasi-steady-state soli-

ton laser operation scenario is the fact that the gain and losses are exactly compensated during one soliton pulse round trip. One can model the effective distributed gain and losses inside the laser cavity by a periodic function such as

$$\Gamma(Z) = \frac{\sin Z}{\Delta^3}; \quad \text{where} \quad \Delta = \sqrt{1 - \delta^2 \sin^2 Z}. \tag{17}$$

The nonlinear Bloch theorem (3), (4) states that the corresponding dispersion function must satisfy

$$D(Z) = -2\frac{\sin Z}{\Delta^3} \exp\left[-\frac{2}{(1 - \delta^2)}\left(1 - \frac{\cos Z}{\Delta}\right)\right]. \tag{18}$$

The main soliton features of the nonlinear solitary Bloch waves predicted analytically were investigated by using direct computer simulations. The nonlinear Bloch waves scenario for the case represented by Eqs. (3), (15), and (16) is shown in Fig. 1. Nontrivial
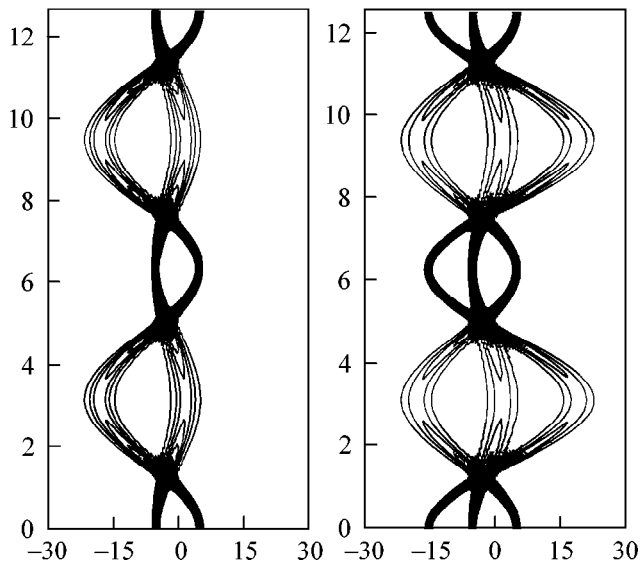
Fig. 3. Elastic interaction between two and three solitary nonlinear Bloch waves in periodic structure given by Eqs. (15) and (16). In the case of temporal solitons, *X* axis corresponds to the dimensionless time in retarded frame; in the case of two-dimensional spatial solitons, *X* axis corresponds to the transverse coordinate. *Y* axis represents the normalized distance of the nonlinear Bloch waves propagation.

(nonzero) initial velocities in the comoving frame of reference are given via the Galileian transformation of Eq. (3). A periodic "snake" effect arises in the space–time domain for solitons with nonzero initial velocities in the retarded frame of reference (see Fig. 2). In Fig. 3, we illustrate typical interactions between nonlinear solitary Bloch waves. An important feature of the solitary nonlinear Bloch waves consists in an elastic character of their interaction, which is shown in Fig. 3.

Recently, the possibility of spatial self-trapping of quasi-periodic waves due to cascaded self-focusing was numerically demonstrated, and the quasi-periodic envelope soliton concept was introduced [8]. The concept of nonlinear Bloch waves, which are localized nonlinear waves in periodic structures, was proposed for the first time by Haus and Chen [9]. Haus and Chen tried to construct the steady-state solutions for the nonlinear problem as a superposition of the Hermit–Gaussians polynomials of the linear propagation problem

[9]. In this letter, the exact solution of the nonlinear problem is obtained. It is shown that solitary nonlinear Bloch waves exist only under certain conditions, and the nonlinear Bloch theorem is proved.

In summary, we predict a novel class of nonlinear solitary bright and dark Bloch waves in inhomogeneous media characterized by translational symmetry. "Classical" solitonlike features of the nonlinear Bloch waves are confirmed by accurate direct computer simulations. The nonlinear Bloch theorem obtained in this letter is of general physical interest and should be readily verified experimentally in periodic nonlinear and dispersive structures in different branches of physics, where the "universal" NLSE model is applicable. For example, it seems very attractive to use the nonlinear Bloch wave concept in ultrashort pulse photonics applications and soliton laser design [10]. The best soliton laser performance is obtained when there is a sign-reversal periodic dispersion and/or nonlinearity inside the laser cavity, according to the nonlinear Bloch theorem.

REFERENCES

1. F. Bloch, Z. Phys. **52**, 555 (1928).
2. L. Brillouin, *Wave Propagation in Periodic Structures* (McGraw Hill, New York, 1946; Inostrannaya Literatura, Moscow, 1959).
3. P. W. Anderson, *Concepts in Solids* (Benjamin, New York, 1964).
4. Melvin Lax, *Symmetry Principles in Solid State and Molecular Physics* (Wiley, New York, 1974).
5. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1974, 5th ed.; Nauka, Moscow, 1978).
6. N. I. Akhiezer, *Elements of Theory of Elliptic Functions* (Gostekhizdat, Moscow, 1948).
7. M. I. Belovolov, E. M. Dianov, V. I. Karpov, *et al.*, Proc. SPIE **963**, 90 (1989).
8. C. B. Clausen, Yu. S. Kivshar, O. Bang, and P. L. Christiansen, Phys. Rev. Lett. **83**, 4740 (1999).
9. H. A. Haus and Y. Chen, J. Opt. Soc. Am. B **16**, 889 (1999).
10. Y. Chen, F. X. Kärtner, U. Morgner, *et al.*, J. Opt. Soc. Am. B **16**, 1999 (1999).

Kolmogorov Spectra in One-Dimensional Weak Turbulence[1]

V. E. Zakharov, O. A. Vasilyev*, and A. I. Dyachenko

*Landau Institute for Theoretical Physics, Russian Academy of Sciences,
Chernogolovka, Moscow region, 142432 Russia*
* e-mail: vasilyev@itp.ac.ru
Received December 16, 2000

In this article, we report the results of our numerical simulation of a one-dimensional modified MMT model, which includes the processes of "one-to-three" wave interactions. We show that this model, with properly chosen parameters, behaves according to the weak-turbulence theory. In particular, it demonstrates the validity of the Kolmogorov spectrum over a wide range of wave numbers. © *2001 MAIK "Nauka/Interperiodica"*.

PACS numbers: 47.27.Eq; 47.35.+i

1. The problem of Kolmogorov spectra is the core of the theory of weak wave turbulence. These spectra appear as exact solutions of stationary kinetic equation for mean squared wave amplitudes [1]. In our opinion, the Kolmogorov weak-turbulence spectra should be used for the theoretical explanation of power-law spectral distributions of energy in ensembles of stochastic nonlinearly interacting waves of any nature. Spectra of such type are observed systematically. The most conspicuous example of this sort is the spectrum $\epsilon_\omega \simeq gv/\omega^4$, which is routinely observed in the systems of wind-driven gravity sea waves.

However, this viewpoint is not shared by everybody. Moreover, the very applicability of the kinetic equation for waves to the real situation is under discussion (see, for instance, [2]). The derivation of the kinetic equation from the initial dynamic equation implies the validity of the assumption of phase randomness, which can be destroyed by the formation of some coherent structures, like solitons or wave collapses. Actually, this criticism has serious foundations. In real situations, the coherent structures are common, but there is no reason for complete abandonment of the weak-turbulence theory. The real life is multicolor, and in many particular situations the coherent structures coexist with weak turbulence, sharing the processes of transport and dissipation of energy and other integrals of motion.

Hence, there is a strong motivation to continue the study of weak turbulence and explore both the case where the coherent structures are important and the case when the influence of such structures is negligible.

2. One of the most promising approaches for the study of weak turbulence is a direct numerical simulation of nonlinear dynamic equations describing wave systems. In many cases, these equations can be effec-tively solved by the use of a spectral code. Of course, numerical simulation of the equation completely relevant to the real physical situation is most preferable. However, a considerable amount of interesting information can be extracted from the solution of simplified, more or less artificial models, which account for the basic features of the real physical equations. Since the weak turbulence is a very general theory, its main statements—applicability of the kinetic equation, existence of the Kolmogorov-type spectra and structures of high-order correlation functions, etc.—can be properly tested with these simple models, for which computer simulation can be carried out more easily.

One-dimensional models are the most attractive ones from this point of view. Even a modest modern computer makes it possible to perform numerical simulation of a system of nonlinear waves including a thousand modes and three decades of scaling. Historically, the first such simulation was accomplished in 1997 by the authors of [2]. They used a model which is called now, after their names, the MMT model. The MMT model is the generalized nonlinear Schrödinger equation conserving not only energy and momentum but also the wave action (number of particles).

Further simulation with the MMT model was performed by the same authors in [3]. Later on, numerical experiments with the MMT model were performed in [4, 5]. The results of both groups basically coincide. The MMT model demonstrates a complicated many-variant behavior that cannot be considered as a certain confirmation of the weak-turbulence theory. In our opinion, this is so because of an interference of the coherent structures, which are present in all versions of the MMT model.

In this article, we report results of our numerical simulation with a modified MMT model, which includes processes "one-to-three" wave interactions not conserving wave action. We will show that this model, with properly chosen parameters, behaves

---

[1] This article was submitted by the authors in English.

according to the weak-turbulence theory. In particular, it demonstrates the validity of the Kolmogorov spectrum in the range of more than two decades.

3. We study the following model:

$$i\left(\frac{\partial \psi}{\partial t} + \gamma_k \psi_k\right) = k^\alpha \psi + a\left\{ \int (|k||k_1||k_2||k_3|)^{\beta/4} \right.$$

$$\times \psi_{k_1}^* \psi_{k_2} \psi_{k_3} \delta_{k+k_1-k_2-k_3} dk_1 dk_2 dk_3$$

$$+ g\int (|k||k_1||k_2||k_3|)^{\beta/4} (\psi_{k_1} \psi_{k_2} \psi_{k_3} \delta_{k-k_1-k_2-k_3} \tag{1}$$

$$\left. + 3\psi_{k_1}\psi_{k_2}^*\psi_{k_3}^* \delta_{k-k_1+k_2+k_3}) dk_1 dk_2 dk_3 \right\}.$$

If $g = 0$, this model becomes the MMT model.

Model (1) can be written as

$$i\left(\frac{\partial \psi}{\partial t} + \gamma_k \psi_k\right) = \frac{\delta H}{\delta \psi_k^*}, \tag{2}$$

where

$$H = H_0 + H_{int}, \quad H_0 = \int |k|^\alpha \psi_k \psi_k^* dk, \tag{3}$$

$$H_{int} = a\left\{ \frac{1}{2}\int (kk_1k_2k_3)^{\beta/4} \psi_k^* \psi_{k_1}^* \psi_{k_2} \psi_{k_3} \delta_{k+k_1-k_2-k_3} \right.$$

$$+ g\int (kk_1k_2k_3)^{\beta/4} [\psi_k^* \psi_{k_1}^* \psi_{k_2}^* \psi_{k_3} + \psi_k \psi_{k_1} \psi_{k_2} \psi_{k_3}^*] \tag{4}$$

$$\left. \times \delta_{k+k_1+k_2-k_3} dk_1 dk_2 dk_3 \right\}.$$

Hamiltonian H describes the following four-wave processes:

(a) scattering obeying the resonant conditions

$$\omega_k + \omega_{k_1} = \omega_{k_2} + \omega_{k_3}, \quad k + k_1 = k_2 + k_3; \tag{5}$$

(b) "one wave-to-three" decay and the reverse process of gluing three waves to one wave, obeying the resonant conditions

$$\omega_k = \omega_{k_1} + \omega_{k_2} + \omega_{k_3}, \quad k = k_1 + k_2 + k_3. \tag{6}$$

Here, $\omega_k = |k|^\alpha$.

We have studied only the case $\alpha > 1$. In this case, resonant conditions (5) have only a trivial solution

$$k_2 = k, \quad k_3 = k_1 \text{ or } k_2 = k_1, \quad k_3 = k, \tag{7}$$

while resonant conditions (6) describe $2 - D$ manifold in space (k, k_1, k_2, k_3). If $a > 0$ and g is small, then Hamiltonian (4) is positively defined. This makes it possible to get rid of any kinds of localized structures.

Under these assumptions, system (1) is described by the kinetic equation

$$\partial n/\partial t + 2\gamma_k n_k + st(n, n, n), \tag{8}$$

where

$$st(n, n, n) = 4\pi a^2 g^2 \left\{ \int (kk_1k_2k_3)^{\beta/2} \right.$$

$$\times (n_{k_1} n_{k_2} n_{k_3} - n_k n_{k_1} n_{k_2} - n_k n_{k_1} n_{k_3} - n_k n_{k_2} n_{k_3})$$

$$\times \delta(k - k_1 - k_2 - k_3)\delta(\omega_k - \omega_{k_1} - \omega_{k_2} - \omega_{k_3}) dk_1 dk_2 dk_3$$

$$+ 3\int (kk_1k_2k_3)^{\beta/2} (n_{k_1} n_{k_2} n_{k_3} + n_k n_{k_1} n_{k_2} \tag{9}$$

$$+ n_k n_{k_1} n_{k_3} - n_k n_{k_2} n_{k_3})\delta(k - k_1 + k_2 + k_3)$$

$$\left. \times \delta(\omega_k - \omega_{k_1} + \omega_{k_2} + \omega_{k_3}) dk_1 dk_2 dk_3 \right\}.$$

By definition,

$$\langle \psi_k \psi_{k'}^* \rangle = n_k \delta_{k-k'}. \tag{10}$$

A stationary equation

$$st(n, n, n) = 0 \tag{11}$$

has, for the proper values of α and β, a power-low solution

$$n_k = \alpha p^{1/3}/k^\lambda, \tag{12}$$

$$\lambda = \frac{2}{3}\beta + 1. \tag{13}$$

This is the Kolmogorov spectrum carrying a constant flux of energy p to the large-k region, and α is the dimensionless Kolmogorov constant.

4. We have performed the numerical simulation of Eq. (1) by the use of the standard spectral code. We set $\alpha = 3/2$, $\beta = 9/4$, $a = 1$ for different values of the dimensionless parameter $g = 0, 0.05, 0.1, 0.15, 0.2$.

Our spectral array included 2048 modes, $-1024 < k < 1023$. The system was pumped at low wave numbers, $\gamma_k = -0.005$ at $5 \le |k| \le 10$. The energy sink at large wave numbers was provided by damping, $\gamma_k = 400(k/512 - 0.5)^2$ at $|k| > 512$.

In all variants of our computations, we observed a growth and stabilization of the total energy H of the wave system. According to the weak-turbulence theory, the stabilization level depends drastically on the g parameter.

Figure 1 clearly demonstrates that $1 \longleftrightarrow 3$ processes play the main role in establishing equilibrium.

Figure 2 displays typical stationary spectra at $g = 0.15$. One can see that, in the range $30 < k < 300$, they can be well approximated by the Kolmogorov exponent $\lambda = 5/2$. A typical value of nonlinearity

$$\epsilon = H_{int}/H$$

is $\epsilon \simeq 0.15$.

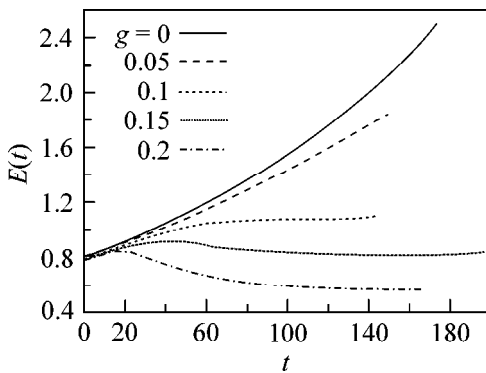In conclusion, we would like to claim that our result is the first clear confirmation of the validity of the

Fig. 1. Total energy of the pumped system versus time for different coefficients of "three-to-one" process.
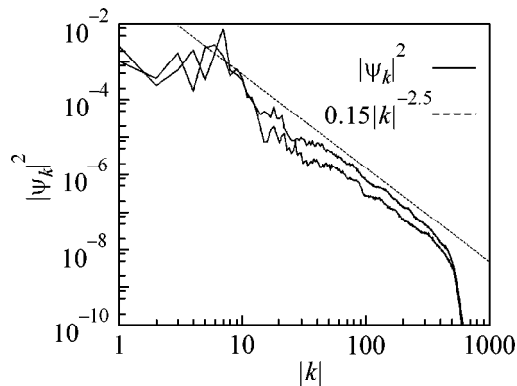

Fig. 2. $|\psi_k|^2$ averaged over time 100. Spectra for positive and negative k are shown.

weak-turbulence theory for the $1 - D$ case. A similar confirmation for the $2 - D$ case was done in work [6]. However, in the present $1 - D$ case, the range of scales where the Kolmogorov spectrum is observed is substantially larger.

REFERENCES

1. V. Zakharov, V. L'vov, and G. Falkovich, *Wave Turbulence* (Springer-Verlag, New York, 1992).

2. A. J. Maida, D. W. McLaughlin, and E. G. Tabak, J. Nonlinear Sci. **6**, 9 (1997).

3. D. Cai, A. J. Maida, D. W. McLaughlin, and E. G. Tabak, Proc. Natl. Acad. Sci. USA **96**, 14216 (1999).

4. V. Zakharov, P. Guyenne, A. Pushkarev, and F. Dias, Preprint of the Centre de Mathematique et Leurs Applcations, Ecole Normal de Cachan #2000-4; Physica D (Amsterdam) (2000) (in press).

5. V. Zakharov, P. Guyenne, A. Pushkarev, and F. Dias, Turbulence d'ondes Dans des Models Unidimensionnels, Corte Rentus (2000) (in press).

6. A. Pushkarev and V. Zakharov, Phys. Rev. Lett. **76**, 3320 (1996).

Anomalous Slowdown of Relaxation in an Ultracold Plasma

A. N. Tkachev and S. I. Yakovlenko

Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 117942 Russia

Received November 20, 2000; in final form, December 15, 2000

Recent experiments by T.C. Killian *et al.* [Phys. Rev. Lett. **83**, 4776 (1999)], in which an ultracold plasma ($N_e \sim 2 \times 10^9$ cm^{-3}, $T_e \sim 0.1$ K, and $T_i \sim 10$ μK) with anomalously long lifetime of ~100 μs was obtained, are explained based on a previously developed theory. The results of computer simulations of the plasma transition into a metastable state and initial heating of electrons up to several K are presented. An expression earlier obtained for the rate of the metastable plasma recombination agrees with the measured anomalously long lifetime. A conclusion is drawn that the previously predicted new physical object—a metastable overcooled plasma—is realized experimentally. © *2001 MAIK "Nauka/Interperiodica".*

PACS numbers: 52.27.Gr; 52.50.Jm; 52.55.Dy

EXPERIMENTAL RESULTS

In [1], an ultracold plasma with unique parameters (the charge density $N_e \sim 2 \times 10^9$ cm^{-3}, electron temperature $T_e \sim 0.1$ K, ion temperature $T_i \sim 10$ μK, and degree of ionization $\alpha \sim 0.1$) was reported to have been obtained. The plasma was produced via two-step ($\lambda_1 \approx 882$ nm and $\lambda_2 \approx 514$ nm) ionization of the xenon $6s[3/2]_2$ metastable state. Xenon metastable atoms were produced in a discharge, slowed down using the Zeeman deceleration technique, accumulated in a magnetic–optical trap, and underwent radiative cooling via the $6s[3/2]_2–6p[5/2]_3$ ($\lambda_1 \approx 882$ nm) transition to a temperature of ~10 μK. The ionizing photon energy ($\lambda_2 \approx 514$ nm) was chosen such that the kinetic energy of an electron created during photoionization would be low, $E/k_B = 0.1$–1000 K.

Such a plasma is strongly nonideal, because the Coulomb coupling parameter $\gamma = e^2/aT_{l,i}$ (the ratio of the mean particle potential energy to the kinetic energy) is large: $\gamma_e = 34$ for electrons and $\gamma_i = 3.4 \times 10^5$ for ions (where e is the electron charge and $a = (4\pi N_e/3)^{-1/3} \sim 5 \times 10^{-4}$ cm is the average distance between charged particles). According to the conventional theory of three-body recombination, such a plasma should instantly be decomposed and heated. The characteristic time of three-body recombination $\tau_{rec}^{(0)} \sim 0.3(m_e^{1/2} T_e^{9/2})/e^{10}N_e^2$ (where e and m_e are the electron charge and mass, respectively) is equal to $\tau_{rec}^{(0)} \sim 5 \times 10^{-16}$ s for $N_e \approx 2 \times 10^9$ cm^{-3} and $T_e \approx 0.1$ K and $\tau_{rec}^{(0)} \sim 2$ ns for $N_e \approx 2 \times 10^8$ cm^{-3} and $T_e \sim 1$ K. However, a plasma lifetime of ~100 μs observed in experiments [1] was many orders of magnitude longer. In [1], it was pointed out that the lifetime of an ultracold plasma was anomalously long, but no explanation for this fact was presented.

THEORETICAL PREDICTION OF A METASTABLE PLASMA

An anomalously slow recombination of an overcooled plasma was predicted in the late 1980s by analyzing the results of simulations based on the fundamental principles of dynamics of a large ensemble of Coulomb particles (see reviews [2–5] and the literature cited therein). To verify this conclusion experimentally, it was proposed to form a bunch of nonideal plasma by ionizing atoms with a laser with a photon energy close to the ionization energy [6, 7]. A concept of "frozen" recombination in a dynamic system was supported in [8]. Below, it is shown that the experimental results [1] can be explained based on the previously developed theory [2–5].

INITIAL STAGE OF RELAXATION

Although, in [1], it was declared that an electron temperature of $T_e \sim 0.1$ K was achieved, the experimental results were presented only for T_e of approximately several K. It is not accidental. The point is that, according to our results, in a time of about one-half of the inverse plasma frequency $t \approx 0.5\omega_L^{-1}$, where $\omega_L = (4\pi e^2 N_e/m_e)^{1/2}$, the electrons are heated due to collective interactions. During this time, the dynamic mixing of phase trajectories of a large ensemble of Coulomb particles occurs. The mixing is characterized by the Lyapunov exponent $L \approx 2.4\omega_L$ [4, 5]. Due to mixing, the coupling parameter decreases to $\gamma \sim 0.4$–0.8.

To simulate experiments [1], we calculated the dynamics of a large ensemble of particles using a technique described in [2, 3]. As before, the electrons and ions were represented by small interpenetrating spheres of diameter $d \approx 0.02 N_e^{-1/3}$ with a uniform surface charge density. The density of charged particles was 2×10^9 cm^{-3}, and the interaction energy of the touching particles was equal to 105 K. The initial conditions, which simulated the photoionization of neutral atoms, were defined as follows. At the initial time, the ions ($n = 512$) were randomly (with a uniform probability density) distributed within a simulation box (a cube with an edge of $(n/N_e)^{1/3}$ with specularly reflecting walls). The initial ion velocities were specified according to the Maxwellian distribution with the temperature $T_i = 10$ μK. Each ion was superimposed by an electron (the initial electron coordinates coincided with the ion coordinates), the electron velocity vectors were uniformly distributed in directions, and the electron kinetic energy was set equal to the ionization energy for a given pair of particles (an "atom"). Then, the Newton equations for $2n$ particles were solved taking into account all the electrostatic interactions, and statistical data were accumulated.

Under the given initial conditions, the stage of electron thermalization is preceded by the escape of electrons from the potential wells at the descending portion of the temperature dependence (see Fig. 1). As multiparticle interactions are involved, the electrons, as in our previous studies, thermalize, and, in a time of $0.5\,\omega_L^{-1} \approx 0.2$ ns, the electron temperature stabilizes due to heating at a level of $T_e \approx 3.3$ K, which corresponds to $\gamma \approx 1$. Then, T_e slowly increases, so that its average value is 4.8 K.

Note that the use of the classical approximation to examine the free particle motion under conditions of experiments [1] is well justified. Even at $T_e \sim 0.1$ K, the quantity $a(m_e T_e)^{1/2}/\hbar$, which characterizes the ratio of the mean interparticle distance to the de Broglie wavelength, is large: $a(m_e T_e)^{1/2}/\hbar \sim 53$ ($a(m_e T_e)^{1/2}/\hbar \sim 370$ at $T_e \sim 5$ K). For the ions and atoms, we have $a(m_{Xe} T_i)^{1/2}/\hbar \sim 260$ at $T_i \sim 10$ μK, where m_{Xe} is the mass of a xenon atom.

Thus, even if laser ionization produces electrons with zero kinetic energy, the electron temperature will stabilize at a value of several K in a time less than one nanosecond. This agrees with the results of [1].

RECOMBINATION MECHANISM

Simulation of the dynamics of a large ensemble of particles [2–5] showed that, in the metastable state, the electron velocity distribution function was Maxwellian; however, the total electron energy distribution function $f(\varepsilon)$ (where ε is the electron energy) differed substantially from the Boltzmann distribution. At neg-
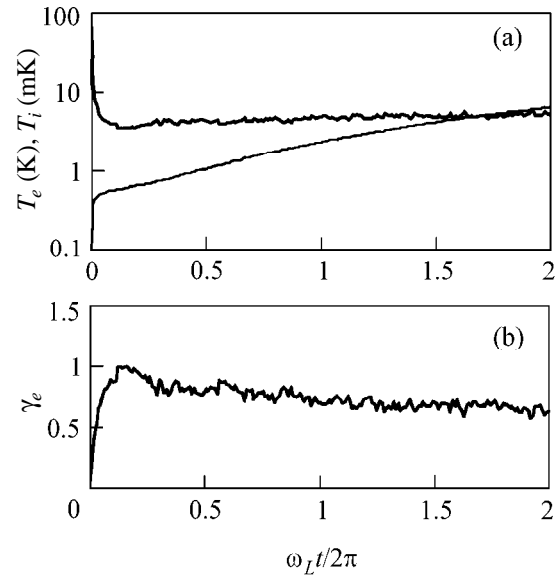

Fig. 1. Time evolution of (a) the ion and electron temperatures (light and heavy curves, respectively) and (b) the coupling parameter $\gamma_e = e^2/(aT_e)$.

ative energies ($\varepsilon < 0$), the distribution function $f(\varepsilon)$ in the metastable state decays exponentially, $f(\varepsilon) \sim \exp(-0.32|\varepsilon|/e^2 N_e^{1/3})$, which drastically differs from the Boltzmann distribution, characterized by exponential growth. A similar result was obtained by us when modeling experiments [1] (see Fig. 2).

Previously, we developed a recombination theory [2, 3] that provided an explanation why the effect of frozen recombination was not observed under ordinary conditions. The theory, which is consistent with the results of computer simulations, is based on the idea that the well-known relations between the kinetic coefficients cannot be used in the region of exponential decay. These relations are based on the detailed equilibrium principle and the assumption that the Boltzmann distribution for the high-lying excited states is equilibrium. Actually, an equilibrium distribution for the high-lying excited states is the above-mentioned distribution with an exponential decay (see [4, 5] for details). However, for the electrons, whose negative energy is high enough,

$$\varepsilon < -\varepsilon_1 \equiv \mathrm{Ry}(e^2 N_e^{1/3}/2\mathrm{Ry})^{2/3},$$
$$\mathrm{Ry} \equiv m_e e^4/2\hbar^2 \approx 13.6 \text{ eV},$$
(1)

a discrete nature of the spectrum becomes of importance, and relaxation due to pair collisions, which can be described by the conventional kinetic models, starts to dominate. Solving the equation for diffusion along the energy axis with allowance for the new relations
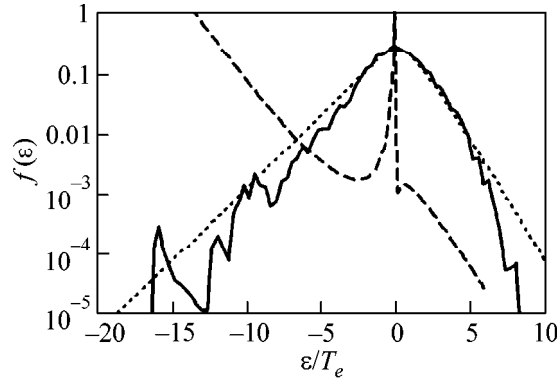
Fig. 2. Electron distribution function over the total energy. Simulation parameters are as follows: the number of charged particles (electrons and ions) is $2n = 1024$ and $N_e = 2 \times 10^9$ cm^{-3}. At the initial instant, the electrons are superposed on the ions and their kinetic energy is equal to the ionization energy. The solid curve is the distribution in a metastable state; it is obtained by averaging over the time interval $\omega_L t = 3.1–12.4$ (see Fig. 1). The dotted and dashed curves are the microfield distribution calculated by analytical formulas from [2–5] and the Boltzmann distribution, respectively (both for $T_e = 3.7$ K).

between the kinetic coefficients leads to the following expression [3]:

$$\tau_{rec} = \tau_{rec}^{(0)} \xi;$$

$$\xi = 1.82 \delta^{5/6} \xi_1 (\varepsilon_1/T_e) \xi_2(N_e) + 6.73 \delta^{7/6}(\xi_2(N_e) - 1). \quad (2)$$

Here, $\tau_{rec}^{(0)}$ is the recombination time predicted by the conventional theory; ξ is a correction factor;

$$\delta = 2e^6 N_e/T_e^3; \quad \xi_1(z) = (e^z/4z^{5/2})$$

$$\times \int_z^\infty dy\, y^{3/2} e^{-y} \sqrt{1 + 6y + 0.75y^2 + (\pi x^3/16)};$$

$$\xi_2(N_e) = \exp(-0.4(\varepsilon_1 - 1.5e^2 N_e^{1/3})/2^{1/3} e^2 N_e^{1/3}).$$

When the electron temperature and density are not too low, expression (2) gives values close to the results of the conventional theory of three-body recombination ($\xi \sim 1$ at $T_e > 0.03$ eV ≈ 350 K and $N_e > 10^{10}$ cm^{-3}). For this reason, we previously concentrated on attempts at creating a strongly overcooled ion–ion plasma, in which quantum effects are of minor importance. It seemed to us that the experimental conditions under which the slowdown of recombination in an electron–ion plasma plays an important role were hard to realize.

However, these hardly achievable conditions were realized in experiments [1]. Indeed, at $T_e = 5$ K and $N_e = 2 \times 10^9$ cm^{-3}, we have $\varepsilon_1 = 56$ K and expression (2) gives a substantial slowdown of recombination compared to the conventional theory: $\xi = 2.4 \times 10^3$. In this case, according to our theory, the characteristic recombination time is $\tau_{rec} = 60$ μs. At $T_e = 5$ K and $N_e = 10^9$ cm^{-3}, we have $\xi = 2.5 \times 10^3$ and $\tau_{rec} = 212$ μs. This is also in agreement with the experimental results [1]. Under these conditions, the radiative recombination, whose characteristic time is about several seconds, can be neglected.

Note that the recombination theories that take into account that the plasma is nonideal [9, 10] also give a recombination time close to $\tau_{rec}^{(0)} \approx 25$ ns for $T_e = 5$ K and $N_e = 2 \times 10^9$ cm^{-3}.

Thus, experiments [1] can naturally be explained by the theory [2–5]. Hence, it is reasonable to assume that a new physical object—a metastable overcooled plasma—was realized in [1]. The initial electron heating occurs due to dynamic relaxation. An anomalous slowdown of relaxation is in quantitative agreement with our theory (up to now, there has been no alternative explanation). The experimental realization of a metastable plasma state is of great importance from the standpoint of the fundamental principles of statistical physics and phase transition theory (see [5] for details). In ion–ion plasmas, the slowdown of relaxation should be even more pronounced [2, 3].

REFERENCES

1. T. C. Killian, S. Kulin, S. D. Bergeson, *et al.*, Phys. Rev. Lett. **83**, 4776 (1999).

2. S. A. Maǐorov, A. N. Tkachev, and S. I. Yakovlenko, Usp. Fiz. Nauk **164**, 298 (1994) [Phys. Usp. **37**, 279 (1994)].

3. S. A. Mayorov, A. N. Tkachev, and S. I. Yakovlenko, Phys. Scr. **51**, 498 (1994).

4. A. N. Tkachev and S. I. Yakovlenko, Izv. Vyssh. Uchebn. Zaved., Fiz. **41**, 47 (1998).

5. S. I. Yakovlenko, Phys. Vib. **6**, 267 (1998).

6. S. I. Yakovlenko, Kvantovaya Élektron. (Moscow) **19**, 5 (1992).

7. S. I. Yakovlenko, Laser Phys. **2**, 196 (1992).

8. B. B. Kadomtsev, *Dynamics and Information* (Redaktsiya Zh. Usp. Fiz. Nauk, Moscow, 1997), IBSN 5-85504-008-9.

9. L. M. Biberman, V. S. Vorob'ev, and I. T. Yakubov, Dokl. Akad. Nauk SSSR **296**, 576 (1987) [Sov. Phys. Dokl. **32**, 752 (1987)].

10. Y. Hahn, Phys. Lett. A **231**, 82 (1997).

Translated by N. Ustinovskiǐ

Homogeneous Electrical Explosion of Tungsten Wire in Vacuum[1]

G. S. Sarkisov[1,2], B. S. Bauer[1], and J. S. De Groot[3]

[1] Department of Physics, University of Nevada, Reno, NV 89557

[2] e-mail: sarkisov@unr.edu.

[3] Department of Applied Science, University of California, Davis, CA 95616

Received December 13, 2000

Experimental results on Joule energy deposition upon initiation of a fast electrical explosion of 16-μm tungsten wire in vacuum at current densities of more than 10^8 A/cm^2 are reported. We have found that explosion with a fast current rise time (~170 A/ns into a short) results in homogeneous and enhanced deposition of electrical energy into the tungsten before surface flashover. The maximum tungsten wire resistivity reaches a value of up to ~185 $\mu\Omega$ cm before surface flashover that significantly exceeds the melting boundary and corresponds to a temperature of ~1 eV. The highest values for light radiation and expansion velocity of wire ~1 km/s were observed for the fast explosion. For the explosion mode with a slower current rise time (~22 A/ns into a short), we observed the existence of an "energy deposition barrier" for tungsten wire. In the slow explosion mode, the current is reconnected to the surface shunting discharge before melting. The maximum tungsten wire resistivity in this case reaches the value of ~120 $\mu\Omega$ cm, which is less than indicative of melting. Also, the energy deposition along the wire is strongly inhomogeneous, and wire is disintegrated into parts. We attribute the early reconnection of the current to the surface discharge for the slow explosion to high electron emission from the wire surface, which starts before melting. © *2001 MAIK "Nauka/Interperiodica".*

PACS numbers: 52.80.Qj

A better understanding of the initial stage in the electrical explosion of refractory metal wires is extremely important for modern *Z*-pinch physics. The impressive result [1, 2] of 1.8 MJ of X rays radiated in 5 ns FWHM, achieved on the *Z*-facility at Sandia National Laboratories, crucially depends on mitigating the Rayleigh–Taylor instability, which degrades plasma compression. The initial perturbations for this instability arise in the plasma shell formed from the exploded wire array. Increased peak X-ray power correlates with decreased wire gap, presumably because this leads to the formation of a smoother plasma shell. Similarly, the dynamics of single-wire *Z*-pinches depends on how energy is initially absorbed by the load.

It is assumed [3] that surface impurities, such as absorbed gases and hydrocarbons, play an important role in energy deposition processes. It is supposed that heated fast-vaporizing impurities create a gas shell around the wire and, after field flashover, the current is switched from the wire to the ionized gas shell. The effectiveness of Joule energy deposition decreases after flashover. The main strategy to avoid the influence of the early surface flashover is preliminary heating of the wire in vacuum for surface degasation. Preheating the wire has been shown [4] to increase the mass and uniformity of the plasma surrounding the cold core. Nevertheless, the maximum wire resistivity in experiments [4] did not reach the level of tungsten melting, and the deposited energy was less than the energy needed for vaporization. The other approaches to increase energy deposition before surface flashover were dielectric coating of the wire [5], explosion of the wire in oil [6], in deionized water [7], and in gas [8]. It has been shown that electrical explosion of tungsten wire in a dense medium and with a coating results in enhanced energy deposition before surface flashover.

We have found a new approach to significantly enhance the efficiency and homogeneity of the electrical explosion of pure tungsten wire in vacuum. In our experiments, we find, for the first time, a connection between the rate of energy deposition and the absolute value and homogeneity of deposited energy into the tungsten wire before surface flashover. For an energy deposition rate of ~0.1 (eV/atom)/ns, we observe an "energy deposition barrier," when surface flashover reconnects current from the wire to the surface before melting. In this case, the wire becomes mainly disintegrated into parts. Parts of the wire that absorbed some energy expand with a velocity of ~0.1–0.2 km/s. For a
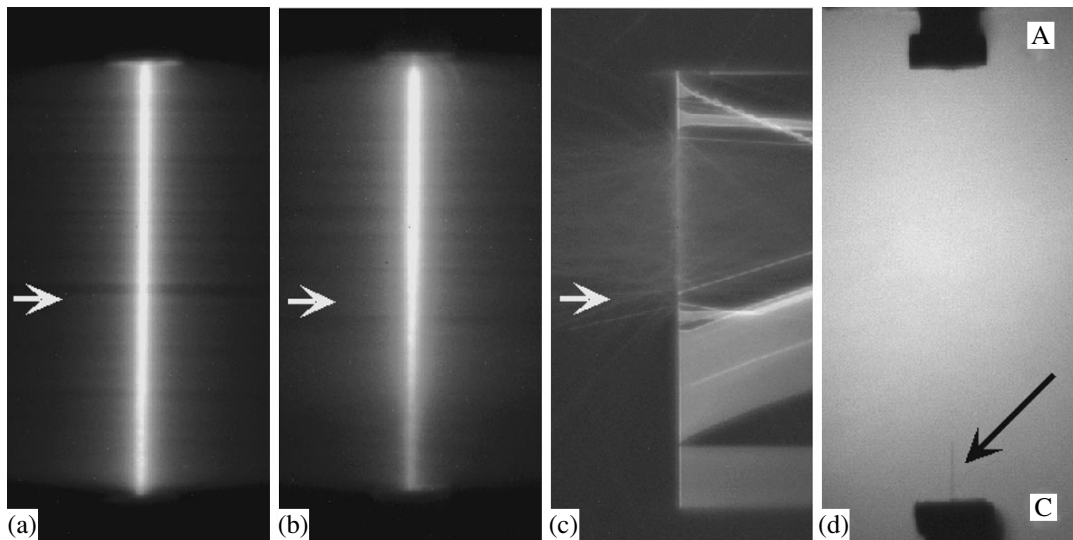

Fig. 1. Time-integrated CCD images of a 16-μm diameter 2-cm long tungsten wire explosion: (a) fast "cylindrical" explosion (shot 0626-08); (b) fast "conical" explosion (shot 0626-11); (c) slow explosion with peripheral vacuum arcs (shot 0801-06); (d) the cathode–anode gap after a slow explosion with unvaporized wire (black arrow). The white arrows show the cross section imaged to the streak camera to investigate the velocity of the wire expansion (Fig. 2). Positions of the anode (A) and cathode (C) are marked on image (d).

faster energy deposition rate, ~0.7 (eV/atom)/ns, we observe a qualitative and quantitative breakthrough in the energy deposition processes. Tungsten wire was exploded homogeneously, and the deposited energy was ~3 times higher. In this case, we observe strong light emission synchronized with surface flashover, and the wire expansion velocity reaches the value ~1 km/s.

A 120-kV Maxwell 40151-B pulse generator with stored energy of 12.6 J (7 nF, 60 kV) provides the current and voltage used to drive the wire explosions. A 50-Ω 9-m coaxial transmission line delivers the positive-voltage pulse from the generator to the wire. The setup produces a fast pulse with a ~170-A/ns current rise into a short circuit (voltage rate ~11 kV/ns in an open circuit) or, with an inductor inserted before the generator output, a slower pulse with a ~22-A/ns current rise into a short circuit (voltage rate ~1 kV/ns into an open circuit). The slower rise is characteristic of the prepulse through individual wires on the *Z*-facility at Sandia National Laboratories. The electrical pulse is applied to a single tungsten wire 2 cm in length and 16 μm in diameter. The wire forms the central conductor of a coaxial line evacuated to a pressure of ~10^{-5} torr to prevent early gas breakdown.

The current flowing through the wire is measured with a 2-GHz bandwidth shunt resistor, and the voltage drop across the wire is measured with a resistive high-voltage probe (100 kV, 1.5 ns rise time). From current and voltage data, the resistive component of the voltage, the load resistance, and the Joule heating of the wire can be determined throughout the wire explosion. The evolution of space-integrated light intensity emitted from the exploding wire is monitored by a PIN

diode with a rise time of less than 1 ns. The light is focused by a f/0.7 lens with a focal length of 50.3 mm. Streak camera shadowgrams of the wire during the explosion are obtained using a high-power diode pulsed laser back lighter (905 nm, 10 W, 200 ns). From these streaked images, the plasma expansion velocity and the starting time of the explosion are determined. Time-integrated visible-light CCD images of wire explosions give information on the two-dimensional structure of the energy deposition. We also analyze the time-integrated emission spectra using a visible imaging spectrometer.

Typical time-integrated images of exploding 16-μm diameter and 2-cm long *W* wire are shown in Fig. 1. There are big differences in the energy deposition structure between the explosions driven by fast and slow pulses. The fast pulse produces a homogeneous "cylindrical" (a) or a "conical" (b) structure. The "conical" explosion (b) is wider at the anode and narrower at the cathode side. The slow pulse results in many peripheral vacuum arcs between the wire and the ground cylinder 2.8 cm away. Part of the wire is left unvaporized (d) after the slow pulse shot (c). For the cylindrical (a) and the conical (b) explosion modes, the region surrounding the wire glows brightly. For the slow pulse case (c), the light emission is over an order of magnitude less and comes mostly from the peripheral vacuum arcs.

The velocity of expansion is dramatically greater for the fast pulse explosion than for the slow pulse case (Fig. 2, data from the shots shown in Fig. 1). The velocity of expansion in the middle of the 16-μm diameter *W* wire is 0.83 km/s for the fast cylindrical explosion
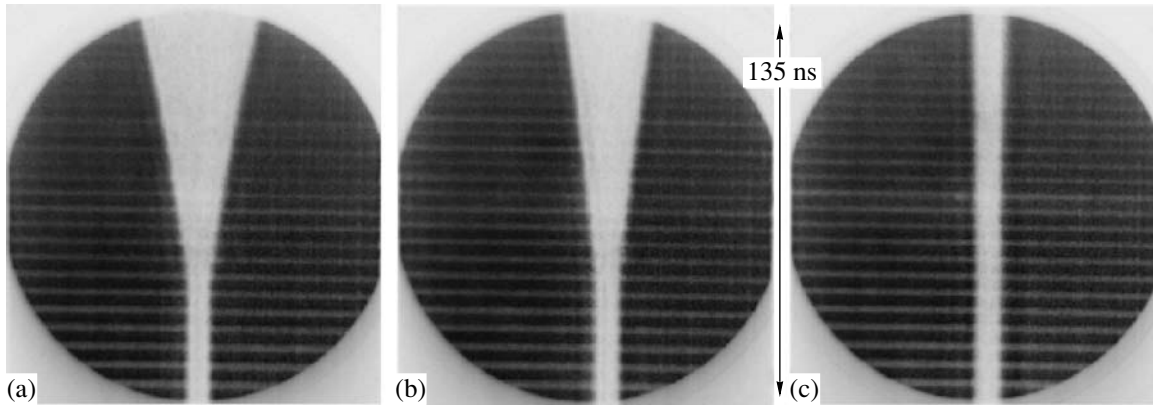
Fig. 2. Radius vs. time (R–T) streaked laser shadowgrams of the expansion in the middle plane of the 16-μm diameter W wire. R–T images (a), (b), and (c) correspond to time-integrated images (a), (b), and (c) in the cross sections marked by white arrows in Fig. 1.

mode (Figs. 2a and 1a), 0.59 km/s for the fast conical explosion mode (Figs. 2b and 1b), and 0.15 km/s for the slow-pulse explosion (Figs. 2c and 1c). In slow-pulse explosion, the wire expands only in the region where the peripheral vacuum arcs are absent. The wire remains unvaporized in the arcing region. The wire starts to expand after the voltage reaches the maximum value, i.e., after surface flashover. The expansion proceeds at a constant speed, indicating minimal heating of the wire after flashover.

Optical spectra for the fast and slow explosion of 16-μm W wire contain a high level of continuum radiation combined with spectral lines of W and some light species.

Temporal evolution of current (1), voltage (2), deposited Joule energy (3), and light emission (4) for fast cylindrical (a) and slow (b) explosion modes of 16-μm W wire is presented in Figs. 3a and 3b. For both types of explosions, we can see the same feature: after reaching the maximum value, the voltage drops fast and then becomes inductive. The rate of specific energy deposition for the fast explosion is seven times higher than for the slow one [~0.7 (eV/atom)/ns for fast explosion and ~0.1 (eV/atom)/ns for slow explosion]. The maximum electrical power for the fast explosion reaches the value of ~61 MW and ~5.5 MW for the slow one. The extremely fast rising (~2 ns initial spike) light intensity for the fast cylindrical explosion coincides in time with the rapid decrease in voltage and rapid increase in current, i.e., with surface flashover. Fast explosions yield the most powerful light emission. The first radiation peak is 4–5 times weaker for wires driven by the slow explosion.

There is a large difference in the long-time behavior of light emission for fast and slow explosion modes. The fast explosion mode gives a strong (3–4 times) increase in the light after the first peak for ~500 ns and a slow drop for ~5–10 μs. For the slow explosion mode, light intensity only drops after the first peak. Interest-

ingly, the long-scale temporal shape of radiation for the fast exploding mode is typical only for W. Other substances give significantly different temporal shapes for the radiation, indicating quite different conditions of the wire core and coronal plasma. For the fast cylindrical mode of explosion, the temporal shape of the radiation depends very strongly and distinctly on the wire material, but not much on the wire diameter.

The dependence of load resistivity (1), current density (2), electrical power (3), and magnetic field pressure (4) on deposited Joule energy for fast cylindrical (c) and slow (d) explosion modes is presented in Figs. 3c and 3d. During the voltage increase, we can attribute the recovered value of resistivity to the property of the wire. When the voltage starts to drop, the current splits between wire and surface flashover and we cannot attribute the recovered value of the resistivity to only the wire or flashover property. The resistivity curves (1) in Figs. 3c and 3d for fast and slow exploding wires coincide for low value of energy deposition. The only difference is the point at which surface flashover occurs (~3 eV/atom for the fast cylindrical mode and ~1.5 eV/atom for the slow explosion mode). For fast explosions, the wire resistivity (1) reaches ~185 $\mu\Omega$ cm, far above the melting boundary [9] corresponding to the saturation plateau in [10]. In contrast, the resistivity with the slowly rising pulse (3) only reaches ~120 $\mu\Omega$ cm, which is below the melting boundary [9]. The maximum current densities before flashover (Figs. 3c, 3d) are substantial, ~250 MA/cm² for fast pulses and ~130 MA/cm² in the slow case. The maximum magnetic field at the wire surface before flashover is ~131 kG and ~67 kG for the fast and slow pulses, respectively. The maximum of the average pressure inside the wire is approximately equal to the magnetic-field pressure on the wire surface and corresponds to ~135 MPa for the fast explosion and ~36 MPa for the slow explosion modes. The magnetic-field pressure in the fast explosion mode is 40% of the critical pressure for W ($p_{cr} = 337$ MPa, [11]) and can suppress boiling at
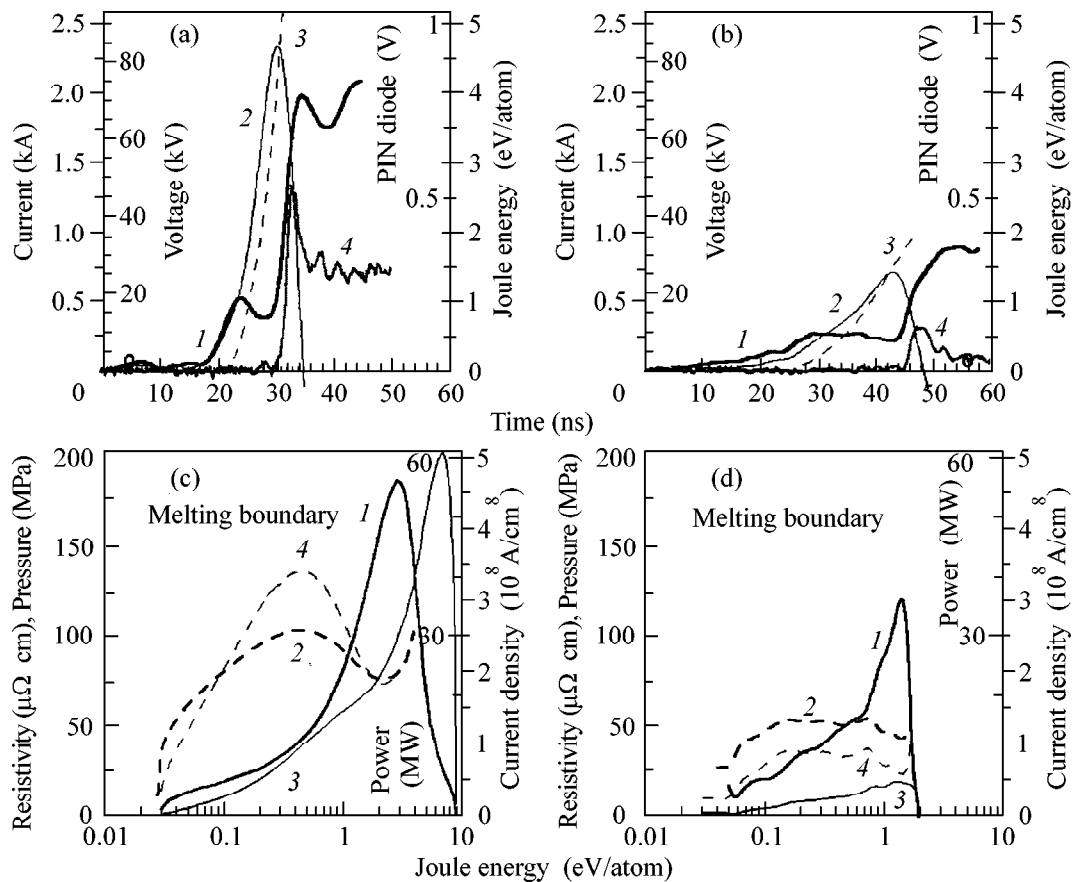
Fig. 3. Temporal evolution of (*1*) current, (*2*) voltage, (*3*) deposited Joule energy, and (*4*) emission light intensity for (a) fast and (b) slow explosions of 16-μm *W* wires. Dependence of (*1*) wire resistivity, (*2*) current density, (*3*) electrical power, and (*4*) magnetic-field pressure on the deposited Joule energy for (c) fast and (d) slow explosions. Magnetic field induction corresponding to current density (*2*) in (c) and (d) has a maximum value of 131 and 67 kG, respectively.

liquid–gas phase transition. In this case, we can assume a homogeneous expansion of the *W* wire.

The results of our experiments suggest the following scenario for the initial stages of the electrical explosion of tungsten wire. For the slow explosion mode, current flows through the wire and increases the temperature for the first ~45 ns. Before melting, the surface flashover starts. Current switches from the wire to surface flashover for a few nanoseconds and practically stops the heating of the substance. In this case, the wire becomes mainly disintegrated on macroparts.

Let us now dwell on flashover reasons. Tungsten is known as a strong electron emitter. It has been found in [6] that, in rapid electrical heating of fine *W* wire in a vacuum, there is an "anomalous electron emission" phenomenon when thermal electron emission exceeds the normal value by ~100 times before melting. Emitted electrons may be a "trigger" for the ionization of vapors surrounding the wire and the creation of a low-resistance plasma shell [12]. Ionization of the vapor shell can be due to direct electron–atom collisions or/and photoionization by X-ray and UV radiation gen-

erated in the collisions of emitted electrons with a high-density wire surface. The strong electron emission from the *W* wire creates an "energy deposition barrier" to effective heating of the substance up to vaporization.

We can support this flashover scenario, because, in the slow pulse case, effective wire resistivity stops to grow at the level of *W* wire melting. Furthermore, we observe a set of peripheral vacuum arcs (Fig. 1c), which can be initiated by electron emission from the wire during the negative period of the electrical pulse (~400 ns after explosion initiation). The high level of continuum radiation in the optical spectra for the fast and slow explosion modes may support the electron emission hypothesis, because it is well known that the interaction of electrons with the wire generates bremsstrahlung radiation. There is another argument for the electron emission hypothesis. Experiments with *W* wire coated with vacuum pump oil demonstrate significantly higher values of deposited energy and expansion velocity than with uncoated wire. In this case, oil can "deactivate" emitted electrons because of high density at the metal–

oil interface. We will describe in detail these experiments in future publications.

We found one of the first descriptions of "wire disintegration regime" in [13] and in recent investigations using high-resolution X-ray back lighting [14]. It has been found that, in the slow explosion mode (~10–20 A/ns), W wire is disintegrated into longitudinal and radial parts, separated into drops [15, 16], and destroyed into microscopic sol with sizes on the order of the electron free path in a metal ~10 nm [17].

Energy deposition into the W wire is significantly improved in the fast implosion regime. In this case, surface flashover starts far after melting. The W wire resistivity reaches a maximum value of ~185 $\mu\Omega$ cm, which is beyond the melting level for W [9] and corresponds to the saturation plateau under normal density [10]. For this value of wire resistivity (at normal density), the temperature before the surface flashover is supposed to be ~1 eV [10]. For the fast explosion mode, the wire is heated at a larger value of current density and magnetic field than for the slow mode. In this case, the larger magnetic insulation of the wire may result in the surface flashover occurring at a temperature that is much higher than the melting temperature. The first narrow peak in light emission, which coincides with time of flashover, can be attributed to the radiation of the exited atoms and recombination of the ions produced due to the ionization of a vapor shell surrounding the wire. Analysis of the role of electron emission in the explosion of the refractory metals wire at current densities ~10^6–10^7 A/cm^2 can be found in [18].

There is no essential difference between the fast cylindrical (Fig. 1a) and the fast conical (Fig. 1b) explosion modes. For the fast conical explosion mode, electron emission may start the flashover at smaller values of absorbed energy and the maximum resistivity does not reach a resistivity plateau [10] but passes through the melting level. For the fast cylindrical explosion mode, flashover starts at higher absorbed energy (close to the vaporization energy) and the maximum resistivity reaches a plateau [10]. The conical deposition energy was observed for many substances with large vaporization energy: W, Mo, Pt, Ti, Ni, and Fe; however, it was never observed for low vaporization energy substances: Au, Cu, Al, and Ag. We will discuss this effect in later publications.

In our experiments, the energy necessary to start wire melting is 1.5 eV/atom, which is ~1.8 times as large as the tabulated value. This agrees with the observation of increased specific heat capacity before melting, discussed in [18]. The high magnetic pressure may also play a certain role in increasing the energy deposition.

Finally, let us summarize the major results of this work. It has been shown that the quality of the electrical explosion of tungsten wire is critically dependent on the rate of energy deposition. The processes vary from electrical disintegrating of the wire into macroscopic

parts (under slow-pulse mode ~20 A/ns) to the homogeneous explosion with a large value of deposited energy (under fast-pulse mode ~170 A/ns). The main barrier to enhanced energy deposition in the slow explosion mode is caused by the initiation of high thermal electron emission before wire melting. Current becomes reconnected to the surface flashover, and heating of the wire stops. This "energy deposition barrier" can be avoided by increasing the wire heating rate. In this case, heating occurs at higher current density, so that a larger magnetic field may restrict the electron emission before melting. In the fast explosion mode, the wire resistivity reaches a value of ~185 $\mu\Omega$ cm, which is a plateau for the normal density of tungsten [10]. In this case, the temperature of the wire before expansion reaches the values of ~1 eV, the energy deposition is homogeneous along the axis, and the expansion velocity reaches a maximum value of ~1 km/s.

The effect of significant enhancement of energy deposition before surface flashover with increasing energy deposition rate was observed for all available pure metal wires (Mo, Pt, Ti, Ni, Fe, Ag, Cu, Al, and Ag); we will discuss the results in future publications.

REFERENCES

1. T. W. Sanford, G. O. Allshouse, B. M. Marder, *et al.*, Phys. Rev. Lett. **77**, 5063 (1996).

2. R. B. Spielman, C. Deeney, G. A. Chandler, *et al.*, Phys. Plasmas **5**, 2105 (1998).

3. F. D. Bennett, in *Progress in High Temperature Physics and Chemistry*, Ed. by C. A. Rouse (Pergamon, Oxford, 1968), Vol. 2, p. 3.

4. S. A. Pikuz, T. A. Shelkovenko, A. R. Minglaev, *et al.*, Phys. Plasmas **6**, 4272 (1999).

5. D. B. Sinars, T. A. Shelkovenko, S. A. Pikuz, *et al.*, Phys. Plasmas **7**, 429 (2000).

6. S. V. Lebedev and S. E. Khaikin, Zh. Éksp. Teor. Fiz. **26**, 723 (1954); **26**, 629 (1954).

7. R. B. Baksht, I. M. Datsko, A. F. Korostelev, *et al.*, Fiz. Plazmy **9**, 1224 (1983) [Sov. J. Plasma Phys. **9**, 706 (1983)].

8. F. H. Webb, H. H. Hilton, P. H. Levine, and A. V. Tollestrup, in *Exploding Wires*, Ed. by W. G. Chace and H. K. More (Plenum, New York, 1962), Vol. 2, p. 37.

9. *Electrical Resistivity Handbook*, Ed. by G. T. Dyos and T. Farrell (Peter Peregrinus, London, 1992).

10. M. P. Desjarlais, Contrib. Plasma Phys. (2001) (in press).

11. M. M. Martynyuk and I. Kasimkhodzhaev, Zh. Fiz. Khim. **48**, 1243 (1974).

12. S. V. Lebedev, High Temp. **8**, 240 (1970).

13. N. N. Sobolev, Zh. Éksp. Teor. Fiz. **17**, 655 (1947); **17**, 987 (1947).

14. S. A. Pikuz, T. A. Shelkovenko, D. B. Sinars, *et al.*, Phys. Rev. Lett. **83**, 4313 (1999).

15. L. Zernow and G. Woffnden, in *Exploding Wires*, Ed. by W. G. Chace and H. K. More (Plenum, New York, 1959), p. 170.

16. J. Nasilowski, in *Exploding Wires*, Ed. by W. G. Chace and H. K. More (Plenum, New York, 1964), Vol. 3, p. 259.

17. F. G. Carioris, B. R. Fish, and G. W. Royster, in *Exploding Wires*, Ed. by W. G. Chace and H. K. More (Plenum, New York, 1962), Vol. 2, p. 299.

18. S. V. Lebedev and A. I. Savvatimskii, Usp. Fiz. Nauk **144** (2), 215 (1984) [Sov. Phys. Usp. **27** (10), 749 (1984)].

Enhanced Superconductivity of the Ti–Zr Alloys
in the High-Pressure BCC Phase*

I. O. Bashkin[1,2], V. G. Tissen[1], M. V. Nefedova[1], A. Schiwek[3],
W. B. Holzapfel[3], and E. G. Ponyatovsky[1]

[1] Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia
[2] e-mail: bashkin@issp.ac.ru
[3] Fachbereich 6 - Physik, Universität-GH Paderborn, D-33095 Paderborn, Germany
Received December 7, 2000

A high-pressure study of the crystal structure and superconductivity of Ti–Zr alloys demonstrates an increase in the ω–β-transition pressure from about 30 to 43–57 GPa when the titanium content in the alloys increases from 0 to 50 at. %. The isobaric values of the BCC β-phase superconducting temperature (at 46 GPa) increase from 5.7 to more than 15 K between 0 and 50 at. % Ti, the latter value being the absolute maximum for BCC d-metal alloys. These data correlate with the earlier assumption of an s–d electron transfer in Zr under pressure. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 64.70.Kb; 74.62.Fj

1. Titanium and zirconium at normal conditions are stable in the hexagonal close-packed structure (the HCP α phase) and transform to the body-centered cubic (BCC) β phase above 1155 and 1136 K, respectively [1]. The hexagonal ω phase becomes stable for both metals under pressures of $P > 2$ GPa at room temperature [1–4]. Recently, Xia *et al.* [5] found that ω-Zr compressed to $P \approx 30$ GPa at room temperature undergoes a transition to a BCC phase which is also called β-phase, by analogy with the low-pressure/high-temperature phase. The hard-sphere packing ratio for the BCC structure is not as high as in the close-packed FCC or HCP structures; therefore Xia *et al.* related this transition to the s–d electron transfer in Zr, which results in a reduction of the atomic volume [5], as discussed in theoretical studies on the relation between the crystal structure and the d-band occupancy in the nonmagnetic d-band metals (see, e.g., [6–11]). Later, Akahama *et al.* [12] measured the pressure dependence of the superconducting temperature, T_c, for Zr and presented new arguments for the s–d electron transfer also in relation to the structural transition and to the similarity in the behavior of Zr under pressure to Group Vb metals. Actually, the T_c values increase under pressure moderately in the stability ranges of the α-Zr and ω-Zr phases, but a sharp jump by several degrees is observed near $P = 30$ GPa [12]. The idea of an electronic transition was used later in order to explain a similar ω–β transition in Hf around 71 GPa [13], as well as an isostructural BCC-to-BCC transition in Zr which was claimed to be seen at 56 GPa [14]. Titanium does not

undergo a transition to the BCC phase at room temperature up to 87 GPa [13].

The hypothesis of a strong s–d electron transfer for the Group IVb metals under pressure opens a door for a new study. In fact, there is a considerable difference between the Group IVb and Vb metals, as well as between the binary Group IVb–Vb and Vb–Vb metal alloys, with respect to their phase diagrams and superconducting properties. For example, the Ti–Zr alloys form homogeneous solid solutions in both liquid and solid states (solid α solutions at low and moderate temperatures and β solutions up to melting) [15]. The superconducting temperatures of pure Ti and Zr, as well as of their HCP α-alloys, are low and show a maximum of $T_c = 1.7$ K at the equiatomic composition, TiZr [16]. The Group Vb metals have a BCC lattice and higher T_c values ($T_c = 5.4$ and 9.26 K for V and Nb, respectively) [15, 16]. The solubility of the Group Vb metals in HCP α-Ti and α-Zr amounts only to a few percent, but the BCC alloys of Group IVb–Vb elements are homogeneous over a much broader concentration range; e.g., the homogeneity ranges of the β phase in the Ti–V or Ti–Nb systems at room temperature are extended from about 20 to 100 at. % V or Nb. The concentration dependence of the superconducting temperature, $T_c(x)$, in the stability range of the BCC Group IVb–Vb alloys has a maximum between 60 and 75 at. % Group Vb metal [15–18].

One can assume from this consideration that $Ti_{1-x}Zr_x$ alloys become similar to the Group IVb–Vb alloys above the high-pressure electronic transition in Zr with the respective consequences; i.e., the BCC solid solutions of titanium in β-Zr extend their stability range
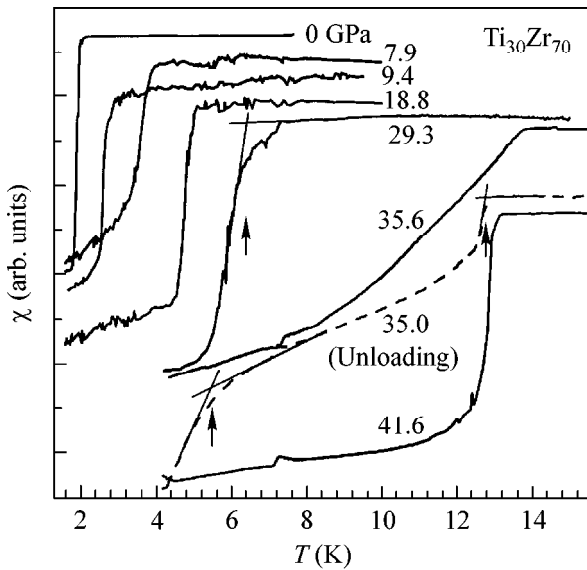

Fig. 1. Representative magnetic susceptibility curves, $\chi(T)$, measured upon heating of a $Ti_{30}Zr_{70}$ alloy at indicated pressures. The arrows explain the determination of the superconducting temperature from the intersection of two tangents to the curve. The curves are measured in a run of stepwise compression, but for the 35.0-GPa isobar (dashed curve) it was measured on decompression.

as the pressure is increased, and the superconducting temperature increases also as the Ti content in β-Zr is increased under isobaric conditions. These assumptions are examined here by T_c measurements for Zr and its alloys with 30 and 50 at. % Ti under pressures up to the transition to the BCC phase, and the α $\longrightarrow$ ω $\longrightarrow$ β structural sequence is confirmed by X-ray diffraction.

2. Experimental. Alloys were prepared from Zr and Ti iodide rods melted in vacuum by electron arc. The purities of molten Zr and Ti were 99.96 and 99.98 at. %, respectively, the interstitial impurities being taken into account. The final alloy compositions were measured by an electron-probe X-ray microanalyzer JXA-5, with the results of $x = 69.3 \pm 0.8$ and $x = 50.4 \pm 0.4$ at. % Zr for the samples used in superconductivity measurements and $x = 71.9 \pm 0.8$ and $x = 50.4 \pm 0.4$ at. % Zr for the structural studies.

The superconducting transitions in the samples were detected as anomalies in the magnetic susceptibility curves, $\chi(T)$, measured with alternating current [19]. The diamond-anvil apparatus made of nonmagnetic materials and the experimental details were described elsewhere [19, 20]. Cooling of the diamond apparatus results in its mechanical relaxation and a marked pressure increase; therefore, pressure determination in these experiments was carried out at room temperature at the end of each experimental cooling/heating cycle. Energy-dispersive X-ray diffraction (EDXD) with synchrotron radiation and diamond-anvil high-pressure cells was performed at HASYLAB (DESY, Hamburg), as described previously in [21, 22].

3. Results. Typical experimental $\chi(T)$ curves for $Ti_{30}Zr_{70}$ measured on heating at constant pressures are represented in Fig. 1. Similar curve shapes were characteristic of Zr and the equiatomic alloy. The jumps due to the superconducting transition are clearly visible in each curve. The jumps in most of the $\chi(T)$ isobars are very steep, indicating a single-phase state of the alloy. There are, however, some pressure intervals where the superconducting $\chi(T)$ anomalies either are extended over several degrees (e.g., the 35.6 GPa isobar in Fig. 1) or consist of two jumps separated by an inclined plateau (the 35.0 GPa decompression isobar in Fig. 1). A strong distortion of the $\chi(T)$ anomalies is typical for pressure intervals, where the ω–β phase transition takes place and the alloy is in a two-phase state. In the pressure intervals of the ω–β transition, distortions of the $\chi(T)$ anomalies are rather small. The superconducting temperature was determined as the intersection point between the steepest tangent to the $\chi(T)$ curve and the linear extension of the high-temperature section of the $\chi(T)$ curve, as shown in Fig. 1.

The corresponding T_c values for the three alloys are plotted versus pressure in Fig. 2 together with the literature data [12, 23, 24] for pure Zr (bottom curve). Figure 2 shows a good agreement between the different data below 30 GPa. Above this pressure, the T_c values from Akahama *et al.* [12] are systematically larger than the present values, probably, due to the difference in the experimental techniques: Akahama *et al.* [12] determined the T_c values from the resistivity measurements, unlike the present $\chi(T)$ data. One can also see from Fig. 2 that the $T_c(P)$ curve in the ω-Zr range is almost a smooth continuation of the $T_c(P)$ curve in the α-Zr range. The ω–β transition occurs in a very narrow pressure interval near 30 GPa, as reported earlier [5, 13]. After the T_c jump from 4.1 to 9.4 K due to the ω–β transition, further pressure increase results in a gradual decrease in T_c.

The alloys have similar $T_c(P)$ dependences in the α- and ω-phase regions: T_c grows with pressure buildup for both phases. Transitions to the ω phase are observed near 10 GPa as a small but distinct decrease in T_c. The slope of the $T_c(P)$ dependence for the ω phase of the equiatomic TiZr alloy is smaller than for the α phase. The process of the ω–β transition in the alloys is different from that in pure Zr. For $Ti_{30}Zr_{70}$, the $\chi(T)$ jump is very broad around 35 GPa on increasing pressure, and two $\chi(T)$ jumps are observed in this range on decompression. This is an indication of an intermediate two-phase state occurring in an extended pressure interval. The transition hysteresis, however, remains rather small. Above the ω–β transition, $T_c(P)$ does not decrease as steeply as in pure Zr. The two-phase nature of the ω–β transition is most clearly seen in the equiatomic TiZr alloy. The broad $\chi(T)$ anomaly around 40 GPa changes to two jumps at $P \geq 43$ GPa, and these two jumps are observed up to the highest pressure of $P = 47$ GPa where T_c for the β phase is 15.5 K.
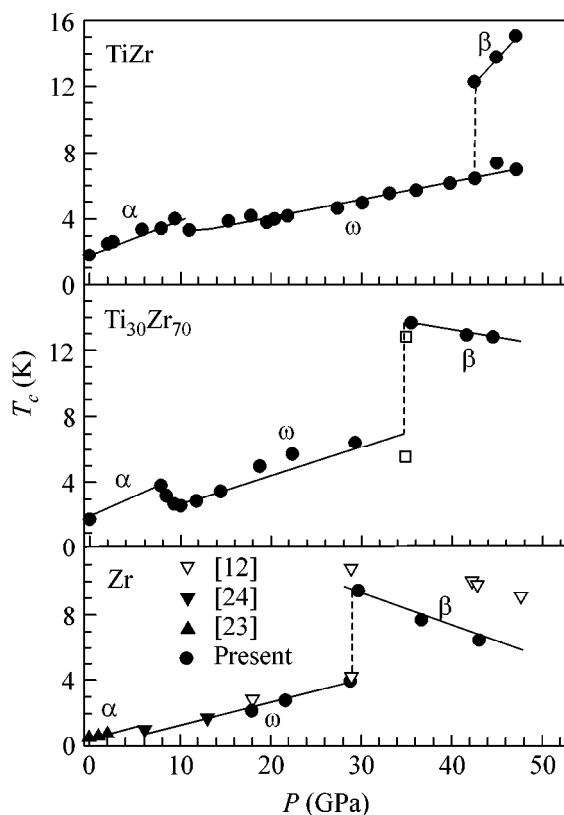
Fig. 2. The pressure effect on the superconducting temperature, $T_c(P)$, for (bottom) Zr, (middle) $Ti_{30}Zr_{70}$, and (top) TiZr, together with literature data for Zr [12, 23, 24]. Open symbols represent the data on decompression of $Ti_{30}Zr_{70}$.
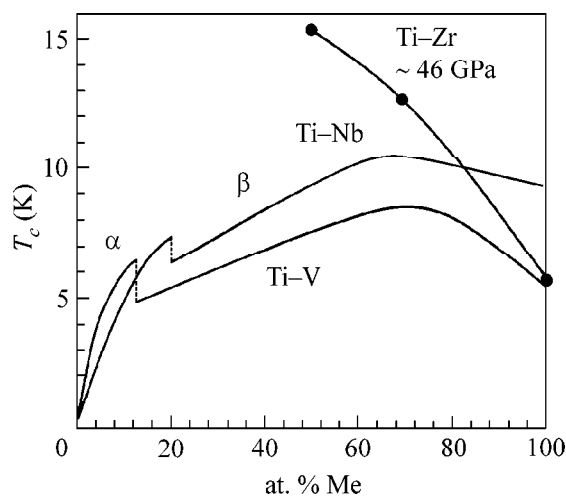

Fig. 3. Effect of the concentration of the second component on the superconducting temperatures of Ti–V and Ti–Nb alloys (reproduced from [16]) and for Ti–Zr alloys (at $P = 46$ GPa). For the Ti–V and Ti–Nb systems, the data in the vicinity of the α/β phase boundary were partly measured on the quenched samples [16].

The *in situ* EDXD measurements confirm for both alloys that the high-pressure T_c jump is related to the formation of the BCC phase. The coexistence interval of the ω and β phases for the TiZr alloy is rather broad: the ω phase is observed up to 56.3 GPa, and the alloy becomes purely β phase at $P = 57.0$ GPa. At 47 GPa, the ω phase is still the dominant phase of the TiZr alloy, with only about 20% of the β phase. So far, T_c seems to increase with pressure, at least until the ω–β transition is completed at 57.0 GPa.

4. Discussion. The present study over a wide pressure range allowed one to determine the $T_c(P)$ dependence for Zr and $Ti_{30}Zr_{70}$ in the pure β phase. For the equiatomic alloy, the T_c value for the pure β phase at the end of the transition should be higher than the value in the two-phase state. For the sake of comparison, Fig. 3 shows the T_c values for the present alloys in the β phase at $P = 46$ GPa and their dependence on the Zr concentration. Figure 3 shows also the earlier ambient-pressure data for the β phase Ti–V and Ti–Nb alloys [16]. From a comparison with these alloys, one could have anticipated, also for the β phase Ti–Zr alloys, only a small initial increase in T_c with increasing Ti content on the Zr-rich side of this diagram. However, the highest observed value of $T_c = 15.5$ K for β TiZr at 47 GPa

is much higher than T_c for pure Zr as well as for any other BCC phase in the transition-metal alloys.

The increase in T_c in the $\omega + \beta$ two-phase region of the TiZr alloy may result from proximity effects in the process of nucleation and growth of the β-phase grains. If this is the case, the T_c values for β-TiZr should further increase until the bulk β grains are well developed.

The T_c values increase upon the ω–β transitions. Thus, the present experimental data support a correlation between the high-pressure structural ω–β transition and the expected s–d electron transfer in Zr. This observation also fits the general assumption of a progressive s–d electron transfer in the earlier transition metals under pressure [6–11]. This correlation can be further elucidated by a more detailed study of $T_c(P)$ for β-TiZr at higher pressures, as well as by further studies closer to the possible maximum in the isobaric $T_c(x)$ curve for the β-$Ti_{1-x}Zr_x$ alloys.

The present determination of the ω–β-transition pressures in the $Ti_{1-x}Zr_x$ alloys at room temperature shows their strong nonlinear increase with the $1 - x$ increase in the range $0 \le 1 - x \le 0.5$ in such a way that the transition in pure Ti may be expected only above 100 GPa. The T_c values also increase with increasing $1 - x$ and become larger than 15 K for TiZr under pressure, which is the record for BCC transition metals and alloys. The T_c increase due to the ω–β transition is in agreement with the general assumption of a progressive s–d electron transfer in the earlier transition metals under pressure.

REFERENCES

1. E. Yu. Tonkov, in *High Pressure Phase Transformations* (Gordon and Breach, Philadelphia, 1992), Vol. 2.

2. V. A. Zilbershtein, N. P. Chistotina, A. A. Zharov, *et al.*, Fiz. Met. Metalloved. **39**, 445 (1975).

3. F. P. Bundy, in *New Materials and Methods for Investigation of Metals and Alloys* (Moscow, 1966), p. 230.

4. A. Jayaraman, W. Klement, and G. C. Kennedy, Phys. Rev. **131**, 644 (1963).

5. H. Xia, S. J. Duclos, A. L. Ruoff, and Y. K. Vohra, Phys. Rev. Lett. **64**, 204 (1990).

6. I. V. Svechkarev and A. S. Panfilov, Phys. Status Solidi B **63**, 11 (1974).

7. C. Duthie and D. G. Pettifor, Phys. Rev. Lett. **38**, 564 (1977).

8. Y. K. Vohra, S. K. Sikka, and W. B. Holzapfel, J. Phys. F **13**, L107 (1983).

9. H. L. Skriver, Phys. Rev. B **31**, 1909 (1985).

10. M. Sigalas, D. A. Papaconstantopoulos, and N. C. Bacalis, Phys. Rev. B **45**, 5777 (1992).

11. R. Ahuja, J. M. Wills, B. Johansson, and O. Eriksson, Phys. Rev. B **48**, 16269 (1993).

12. Y. Akahama, M. Kobayashi, and H. Kawamura, J. Phys. Soc. Jpn. **59**, 3843 (1990).

13. H. Xia, G. Parthasarathy, H. Luo, *et al.*, Phys. Rev. B **42**, 6736 (1990).

14. Y. Akahama, M. Kobayashi, and H. Kawamura, J. Phys. Soc. Jpn. **60**, 3211 (1991).

15. M. Hansen and K. Anderko, in *Constitution of Binary Alloys* (McGraw-Hill, New York, 1958; Metallurgizdat, Moscow, 1962), Vol. 2.

16. E. M. Savitskii, Yu. V. Efimov, N. D. Kozlov, *et al.*, *Superconducting Materials* (Metallurgizdat, Moscow, 1976).

17. G. Gladstone, M. A. Jensen, and J. R. Schrieffer, in *Superconductivity*, Ed. by R. D. Parks (Marcel Dekker, New York, 1969).

18. S. V. Vonsovskii, Yu. A. Izyumov, and E. Z. Kurmaev, *Superconductivity of Transition Metals, Their Alloys and Compounds* (Nauka, Moscow, 1977).

19. V. G. Tissen, E. G. Ponyatovskii, M. V. Nefedova, *et al.*, J. Phys.: Condens. Matter **8**, 3069 (1996).

20. I. O. Bashkin, M. V. Nefedova, and V. G. Tissen, Fiz. Tverd. Tela (St. Petersburg) **42**, 12 (2000) [Phys. Solid State **42**, 11 (2000)].

21. K. Syassen and W. B. Holzapfel, Europhys. Conf. Abstr. **1A**, 75 (1975).

22. W. B. Holzapfel, in *High Pressure Chemistry*, Ed. by H. Kelm (D. Reidel, Boston, 1978), NATO ASI Ser., Ser. C **41**, 177 (1978).

23. N. B. Brandt and N. I. Ginzburg, Usp. Fiz. Nauk **85**, 202 (1965).

24. A. Eichler and W. Gey, Z. Phys. **251**, 321 (1972).

Nonadiabatic Effects in the Phonon Spectrum of Metal Compounds

A. I. Morosov

Moscow State Institute of Radio Engineering, Electronics, and Automation (Technical University),
pr. Vernadskogo 78, Moscow, 117454 Russia

e-mail: morosov@eot.mirea.ru

Received December 7, 2000

It is shown that nonadiabatic corrections to the dispersion law of optical phonons in the region of small wave vectors in the case of branches, for which the vibration with a zero wave vector is not accompanied by the appearance of a dipole moment in the ionic lattice, are significant for all possible directions of the wave vector. If a dipole moment arises, nonadiabatic corrections reach a noticeable value only for the wave-vector directions that are almost perpendicular to the direction of the dipole moment. © *2001 MAIK "Nauka/Interperiodica".*

PACS number: 63.20.-e

The adiabatic approximation, which suggests that the electron density distribution corresponds to an instantaneous configuration of the ionic lattice, is commonly used to calculate phonon spectra. It was shown [1] that the nonadiabatic corrections in the absence of nesting in the electronic system of the crystal can become significant at $\omega_p(\mathbf{q}) \gtrsim q v_F$ ($\mathbf{q}$ is the wave vector, $\omega_p(\mathbf{q})$ is the adiabatic dispersion law for the pth phonon branch, and v_F is the Fermi velocity of electrons), i.e., for the optical phonon branches in the region of small wave vectors. The nonadiabatic corrections are small for $q \approx q_B$, where q_B is the Brillouin wave vector, because they are proportional to the $(m/M)^{1/2}$ ratio (m is the electron mass and M is the ion mass).

It is in the region of wave vectors $q \lesssim \omega_p(\mathbf{q})/v_F \ll q_B$ that the corrections to the bare vertex of electron–phonon interaction become significant [2].

These corrections were taken into account in [3, 4] under the assumption that the electron screening effects are insignificant. At the same time, it was demonstrated in [5] that the screening effects can almost fully compensate for the nonadiabatic corrections.

This article is devoted to examining the screening process and to finding the phonon modes for which the nonadiabatic corrections are not compensated.

The electron–phonon interaction Hamiltonian has the form

$$H_{e-ph} = \sum_{n, n', \mathbf{k}, \mathbf{k'}} \Gamma_{n, n', \mathbf{k}, \mathbf{k'}} \hat{a}_{n'\mathbf{k'}}^{+} \hat{a}_{n, \mathbf{k}}$$
$$\times [\hat{b}(\mathbf{k'} - \mathbf{k}) + \hat{b}^{+}(\mathbf{k} - \mathbf{k'})], \quad (1)$$

where $\hat{a}_{n, \mathbf{k}}^{+}$ and $\hat{a}_{n, \mathbf{k}}$ are the second quantization operators of electrons belonging to the nth band; $\mathbf{k}$ is the wave vector; and $\hat{b}^{+}(\mathbf{q})$, $\hat{b}(\mathbf{q})$ are the second quantization operators of phonons. For simplicity, we will restrict our consideration to one optical phonon branch with the adiabatic dispersion law $\omega_0(\mathbf{q})$. The unscreened matrix element of the electron–phonon interaction has the form

$$\Gamma_{n, n', \mathbf{k}, \mathbf{k'}} = -\sum_{s=1}^{r} \left(\frac{\hbar N}{2 M_s \omega_0(\mathbf{k'} - \mathbf{k})} \right)^{1/2}$$
$$\times \int \psi_{n', \mathbf{k'}}^{*}(\mathbf{r})(\nabla W_s(\mathbf{r}), \mathbf{e}_s(\mathbf{k'} - \mathbf{k}))\psi_{n, \mathbf{k}}(\mathbf{r}) d^3\mathbf{r}, \quad (2)$$

where the summation is over all ions in the unit cell, N is the number of unit cells in the crystal, $\psi_{n, \mathbf{k}}(\mathbf{r})$ is the Bloch function of an electron of the nth band, $W_s(\mathbf{r})$ is the potential energy of electron interaction with the sth ion of mass M_s and charge Z_s in the equilibrium position, and $\mathbf{e}_s(\mathbf{q})$ is the polarization vector of the corresponding ion for a given vibrational mode.

Consider the diagonal matrix element ($n = n'$). The major contribution to the integral on the right-hand side of Eq. (2) comes from distances $r \approx |\mathbf{k'} - \mathbf{k}|^{-1}$ large as compared to the interatomic distance d. It can be assumed that

$$\psi_{n, \mathbf{k}}(\mathbf{r}) \equiv u_{n, \mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k'r}) \approx u_{n, \mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k'r}), \quad (3)$$

and the Bloch function can be replaced by plane waves. Then, in the rigid-ion approximation, $\Gamma_{n, n, \mathbf{k}, \mathbf{k'}} = \Gamma(\mathbf{k'} - \mathbf{k})$;

i.e., this quantity depends only on the difference $\mathbf{k}' - \mathbf{k}$ and equals

$$\Gamma(\mathbf{q}) = -i\left(\frac{\hbar N}{2\omega_0(\mathbf{q})}\right)^{1/2}\frac{4\pi e^2}{q^2}(\mathbf{q}, \mathbf{A}), \qquad (4)$$

where

$$\mathbf{A} = \sum_{s=1}^{r}\frac{Z_s}{M_s^{1/2}}\mathbf{e}_s(0). \qquad (5)$$

Two cases are possible in the region $q \ll q_B$:

(a) If $\mathbf{A} \neq 0$ and the angle between $\mathbf{q}$ and $\mathbf{A}$ is not too small, $\Gamma(\mathbf{q}) \propto q^{-1}$.

(b) If $\mathbf{A} = 0$ or the cosine of the angle φ between $\mathbf{q}$ and $\mathbf{A}$ is small ($\cos\varphi \lesssim q/q_B$), the next terms of the expansion of $u_{n,\,\mathbf{k}'}(\mathbf{r})$ in terms of q in the vicinity of $\mathbf{k}' = \mathbf{k}$ must be taken into account. In this case, the $\Gamma_{n,\,n,\,\mathbf{k},\,\mathbf{k}'}$ function no longer has a singularity at $\mathbf{k}' \longrightarrow \mathbf{k}$, depends on $\mathbf{k}'$ and $\mathbf{k}$ separately, and is of the same order of magnitude as the nondiagonal ($n \neq n'$) matrix elements, for which the distances $r \approx d$ make the major contribution when taking the integral in Eq. (2).

The Coulomb matrix element of electron–electron interaction can be considered in a similar way. Because of the scalar nature of a Coulomb vertex, divergence always takes place at small transferred momenta if both interacting electrons (from n and n' bands, respectively) remain in the initial bands

$$V_{n,\,n'}(\mathbf{q}) = 4\pi e^2/q^2. \qquad (6)$$

Therefore, one may neglect the other nonsingular contributions to screening in the region of small $\mathbf{q}$ and restrict the consideration to the random phase approximation.

Let us consider the renormalization of the phonon Green's function $D(\mathbf{q}, \omega)$ in the presence of nonadiabatic corrections

$$D^{-1}(\mathbf{q}, \omega) = \frac{\omega^2 - \omega_0^2(\mathbf{q})}{2\omega_0(\mathbf{q})} - \Delta\Pi(\mathbf{q}, \omega), \qquad (7)$$

where

$$\Delta\Pi(\mathbf{q}, \omega) = \Pi(\mathbf{q}, \omega) - \Pi(\mathbf{q}, 0), \qquad (8)$$

while the polarization operator $\Pi(\mathbf{q}, \omega)$, with regard to screening, is given, similarly to [5], by the following series:

$$\Pi(\mathbf{q}, \omega) = \sim\!\!\oslash\!\!\sim + \sim\!\!\oslash\!\!\text{----}\!\!\oslash\!\!\sim \qquad (9)$$
$$+ \sim\!\!\oslash\!\!\text{----}\!\!\oslash\!\!\text{----}\!\!\oslash\!\!\sim + \dots,$$

where a vertex with a wavy line corresponds to $\Gamma_{n,\,n',\,\mathbf{k},\,\mathbf{k}'}$, the dashed line corresponds to the Coulomb interaction, and the shaded loop is formed by two electron Green's functions

$$G_n(\mathbf{k}, \varepsilon) = [\varepsilon - \varepsilon_n(\mathbf{k}) + \mu - \Sigma_n(\mathbf{k}, \varepsilon)]^{-1}, \qquad (10)$$

where $\varepsilon_n(\mathbf{k})$ is the dispersion law for electrons in the nth band, μ is their chemical potential, and $\Sigma_n(\mathbf{k}, \varepsilon)$ is the self-energy part. The loop contains all ladder-type corrections caused by the electron–phonon interaction

$$\oslash = \bigcirc + \text{(loop)} + \text{(loop)} + \dots. \qquad (11)$$

The summation in Eq. (9) goes over all indices of bands n.

In the case (a), the contribution to $\Pi(\mathbf{q}, \omega)$ from the terms diagonal in n and relating to partially occupied bands takes the following form after making allowance for the screening effects:

$$\Pi(\mathbf{q}, \omega) = -\frac{|\Gamma(\mathbf{q})|^2 P(\mathbf{q}, \omega)}{1 + 4\pi e^2 P(\mathbf{q}, \omega)/q^2}, \qquad (12)$$

where $P(\mathbf{q}, \omega)$ is given by series (11).

The second term in the denominator of Eq. (12) by far exceeds the first one in the region of small wave vectors. Therefore, to the corrections of order m/M, $P(\mathbf{q}, \omega)$ cancels at $q \lesssim \omega_0/v_F$, and the corresponding contribution to $\Delta\Pi(\mathbf{q}, \omega)$ equals zero. The ladder corrections to the vertex are small for the terms nondiagonal in n, and the same is true for the nonadiabatic corrections [which are of order $(m/M)^{1/2}$]. Thus, in the case (a), the nonadiabatic corrections to the adiabatic dispersion law of optical phonons in the region of small wave vectors are of the same order of smallness, $(m/M)^{1/2}$, as at $q \approx q_B$, and the dispersion is small in the region $q \lesssim \omega_0/v_F$. It is this case that was considered in [5].

In the case (b), the contribution of the terms diagonal in n is no longer expressed by Eq. (12). The point is that $\Gamma_{n,\,n,\,\mathbf{k},\,\mathbf{k}'}$ depends on the angle between $\mathbf{k}$ and $\mathbf{q} = \mathbf{k}' - \mathbf{k}$. Therefore, the integral over $\mathbf{k}$ in a loop containing one Coulomb and one phonon vertex gives a zero result to the leading order in q [6]. At the same time, the integral in a loop containing two phonon vertices is nonzero. Thus, in the case (b), the screening effects are unimportant, the nonadiabatic corrections to the terms nondiagonal in n are significant, and the contribution from the nondiagonal terms is the same as in the case (a). Finally, for the case (b), the nonadiabatic corrections to the adiabatic dispersion law of optical phonons in the region $q \lesssim \omega_0/v_F$ are on the order of ω_0 and must be taken into account in the calculation of phonon spectra. It is precisely the case (b) to which the results of calculations in [3, 4] relate.

Because the displacement of the sth atom in the unit cell for a given phonon mode is proportional to $\mathbf{e}_s/M_s^{1/2}$, the vector $\mathbf{A}$ [Eq. (5)] is proportional to the ionic dipole moment of the unit cell, which arises for an optical vibration with $\mathbf{q} = 0$. If $\mathbf{A} = 0$ for a given optical phonon branch (that is, the dipole moment does not arise), the nonadiabatic corrections are significant for all $q \lesssim \omega_0/v_F$. If $\mathbf{A} \neq 0$ for a given optical branch (a dipole moment arises), the nonadiabatic corrections are signif-

icant only for $\mathbf{q}$ directions that are virtually perpendicular to the dipole moment vector, namely, for the stripe

$$(\mathbf{q}, \mathbf{A})/qA \lesssim q/q_B. \tag{13}$$

In these cases, a notable dispersion due to nonadiabatic corrections must occur in the region of wave vectors $q \lesssim \omega_0/v_F$, which can be observed experimentally. Because the adiabatic approximation overestimates the screening effect, the nonadiabatic corrections must lead to a rise in phonon frequency. This behavior of the phonon spectrum was observed for osmium [7] by Raman scattering spectroscopy.

If the nonadiabatic corrections for a given $\mathbf{q}$ are significant for several optical branches, their mixing must be taken into account. This mixing is accomplished by the polarization operator nondiagonal in the phonon branch number, i.e., having vertices relating to different phonon branches. The hybridization with the other phonon branches for which the adiabatic approximation is adequate is insignificant.

Of interest is an experimental investigation of the angular dependence of dispersion in the region of small wave vectors for phonon modes with $\mathbf{A} \neq 0$. The investigation of centrosymmetric metal compounds, for which these modes are Raman-inactive, can be carried out by hyper-Raman or neutron scattering spectroscopy.

REFERENCES

1. S. Englesberg and J. R. Schriffer, Phys. Rev. **131**, 993 (1963).
2. A. B. Migdal, Zh. Éksp. Teor. Fiz. **34**, 1438 (1958) [Sov. Phys. JETP **7**, 996 (1958)].
3. I. P. Ipatova and A. V. Subashchiev, Zh. Éksp. Teor. Fiz. **66**, 722 (1974) [Sov. Phys. JETP **39**, 349 (1974)].
4. E. G. Maksimov and S. V. Shulga, Solid State Commun. **97**, 553 (1996).
5. B. T. Geĭlikman and A. I. Morozov, Fiz. Tverd. Tela (Leningrad) **18**, 3331 (1976) [Sov. Phys. Solid State **18**, 1941 (1976)].
6. L. A. Fal'kovskiĭ, Zh. Éksp. Teor. Fiz. **103**, 666 (1993) [JETP **76**, 331 (1993)].
7. Yu. S. Ponosov, G. A. Bolotin, C. Thomson, and M. Cardona, Phys. Status Solidi B **208**, 257 (1998).

Translated by A. Bagatur'yants

Coulomb Repulsion-Induced Hyperbolic Pairing as a Mechanism of High-T_c Superconductivity

V. I. Belyavskiĭ[1] and Yu. V. Kopaev[2]

[1] *Voronezh State Pedagogical University, Voronezh, 394043 Russia*
[2] *Lebedev Physical Institute, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 117924 Russia*
Received December 8, 2000

The hyperbolic metric of the dispersion law (the effective mass tensor components of carriers are opposite in sign) in the vicinity of the Fermi contour in high-T_c superconducting cuprates in the case of repulsive interaction gives rise to a superconducting state characterized by the condensate of pairs with a large total momentum (hyperbolic pairing). The gain in the energy of the superconducting state over the normal state is due to the fact that a change in the kinetic energy of pairs (because of the negative light component of the effective mass) dominates over the change in the potential energy (corresponding to energy loss). The shift of the chemical potential upon the transition to the superconducting phase is substantial in this case. With increasing repulsive interaction, the superconducting gap Δ_K increases and the resulting gain in energy changes to an energy loss at a certain critical value of the repulsive potential. The low temperature T_c of the superconducting transition and the large value of Δ_K in this region of potential values are the reasons for the high value of the $2\Delta_K/T_c$ ratio and for the developed quantum fluctuations that are observed in underdoped cuprate superconductors. © 2001 MAIK "Nauka/Interperiodica".

PACS number: 74.20.-z

When quasiparticles in a Fermi system are paired into pairs with a nonzero quasimomentum K, the quasimomentum of the relative motion of a pair belongs to a certain region Ξ_K inside the Brillouin zone [1, 2]. The size and the shape of this region depend on K and are determined by the condition that the quasimomenta of two quasiparticles forming a pair must be either outside or inside the region bounded by the Fermi surface.

The electronic states in high-T_c superconducting compounds are of quasi-2D character, and the structure of energy bands is such that the equienergy lines in the vicinity of the Fermi level contain long, almost rectilinear sections, which provides almost ideal Fermi-contour nesting [3]. Thus, if pairs with $K \ll k_F$ rather than $K \approx 2k_F$ [4, 5] are considered, where k_F is the Fermi momentum for a given direction, the Ξ_K region has a form of a long narrow stripe, as is shown schematically in Fig. 1 for the case when K is aligned with the [010] axis. The k_1 axis divides the Ξ_K region into two parts. If in one of the parts (the upper part in Fig. 1a) the quasimomentum of the relative motion of a hole pair equals k and the spin of a hole with the quasimomentum $k_+ = K/2 + k$ is, by convention, oriented up ($\uparrow$), then in the other (lower) part the quasimomentum of the relative motion equals $-k$ and the spin of a hole with the quasimomentum $k_- = K/2 - k$ at singlet pairing has the orientation $\downarrow$. Inside Ξ_K disposed in the vicinity of a saddle point of the dispersion law, holes have a rather weak dispersion (extended flat-band saddle point [3]), and the (negative) effective mass along the k_2 axis is significantly smaller in magnitude than the effective mass along the k_1 axis (the corresponding Fermi velocities v_{F1} and v_{F2} along k_1 and k_2 differ from each other by more than an order of magnitude, $v_{F1}/v_{F2} < 0.1$ [6]).

The singularity of the scattering amplitude for the relative motion of two holes [2] subject to Coulomb repulsion (Fig. 2) points to the possibility of pairing $(k_+ \uparrow; k_- \downarrow)$. Pairs with a positive decay correspond to quasistationary states that can be responsible for the formation of a pseudogap [1]. The occurrence of states with a negative decay [2] reflects the tendency toward a rearrangement of the ground state of the system with the formation of a superconducting pair condensate.

Consider a Ξ_K region corresponding to hole energies lower than the Fermi energy (Fig. 1). Let us write the Hamiltonian describing the relative motion of hole pairs with quasimomentum K in the form

$$\hat{H}_K = \sum_{k \in \Xi_K} \left\{ \varepsilon\left(\frac{K}{2} + k\right) \hat{a}^+_{\frac{K}{2}+k\uparrow} \hat{a}_{\frac{K}{2}+k\uparrow} \right.$$
$$\left. + \varepsilon\left(\frac{K}{2} - k\right) \hat{a}^+_{\frac{K}{2}-k\downarrow} \hat{a}_{\frac{K}{2}-k\downarrow} \right\} \quad (1)$$
$$+ \sum_{k, k' \in \Xi_K} U(k-k') \hat{a}^+_{\frac{K}{2}+k\uparrow} \hat{a}^+_{\frac{K}{2}-k\downarrow} \hat{a}_{\frac{K}{2}-k'\downarrow} \hat{a}_{\frac{K}{2}+k'\uparrow}.$$

Here, $\varepsilon(k)$ is the energy of a hole with quasimomentum k, $\hat{a}^+_{k\alpha}$ ($\hat{a}_{k\alpha}$) is the creation (annihilation) operator of a
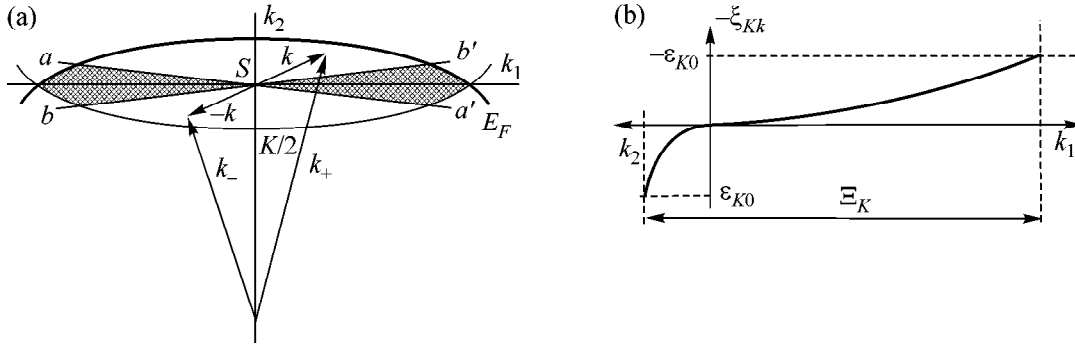
Fig. 1. (a) Domain of definition of the quasimomentum of the relative motion of a pair with quasimomentum K. The section of the hole Fermi contour is distinguished by a thick line. The part of the Ξ_K region with a negative excitation energy ($\xi_{Kk} < 0$) is shaded. In the unshaded area, $\xi_{Kk} > 0$. The kinetic energy of a pair goes to zero on lines a–a' and b–b'. The point S is a saddle point of the dispersion law of a pair. (b) Illustration of the dispersion law of a pair for directions k_1 and k_2.

hole with quasimomentum k and spin projection α, $U(k)$ is the Fourier transform of the energy of correlation interaction between two holes, and the summation over quasimomenta is restricted to the Ξ_K region. Choosing the Hamiltonian in the form of Eq. (1) corresponds to the Bardeen–Cooper–Schrieffer (BCS) approximation in the sense that only the interactions directly leading to the pairing of holes with total quasimomentum $K \approx 2k_F$ are taken into account. As usual, it is assumed that the interactions not included in Eq. (1) exert no significant direct effect on the condensation energy. The average number of holes participating in pairing is evidently constant. Therefore, the corresponding condition for the operator

$$\hat{N}_K = \sum_{k \in \Xi_K} \left(\hat{a}^+_{\frac{K}{2}+k\uparrow} \hat{a}_{\frac{K}{2}+k\downarrow} + \hat{a}^+_{\frac{K}{2}-k\downarrow} \hat{a}_{\frac{K}{2}-k\downarrow} \right) \qquad (2)$$

can be, as usual, taken into account in the diagonalization of Eq. (1) by the Lagrange multiplier method. Denote a Lagrange multiplier by μ_K and introduce new quasiparticle operators by using the Bogoliubov transformation

$$\hat{a}_{\frac{K}{2}+k\uparrow} = u_{Kk}\hat{b}_{k,+1} + v_{Kk}\hat{b}^+_{k,-1},$$

$$\hat{a}^+_{\frac{K}{2}+k\uparrow} = u_{Kk}\hat{b}^+_{k,+1} + v_{Kk}\hat{b}_{k,-1},$$

$$\hat{a}_{\frac{K}{2}-k\downarrow} = u_{Kk}\hat{b}_{k,-1} - v_{Kk}\hat{b}^+_{k,+1}, \qquad (3)$$

$$\hat{a}^+_{\frac{K}{2}-k\downarrow} = u_{Kk}\hat{b}^+_{k,-1} - v_{Kk}\hat{b}^+_{k,+1}.$$

This allows the Hamiltonian in Eq. (1) to be represented, to an accuracy of $\hat{b}^2$, in the form

$$\hat{H}_K = E_{K0} + \hat{H}^{(0)}_K + \hat{H}^{(1)}_K. \qquad (4)$$

The ground-state energy can be written as

$$E_{K0} = -2 \sum_{k \in \Xi_K} \xi_{Kk} v^2_{Kk} + \sum_{k \in \Xi_K} \Delta_{Kk} u_{Kk} v_{Kk}, \qquad (5)$$

where $2\xi_{Kk} = 2\mu_K - \varepsilon(K/2 + k) - \varepsilon(K/2 - k)$ and, by definition,

$$\Delta_{Kk} = \sum_{k \in \Xi_K} U(k - k') u_{Kk'} v_{Kk'}. \qquad (6)$$

The part of the Hamiltonian diagonal in the operators $\hat{b}$ can be represented as

$$\hat{H}^{(0)}_K = \sum_{\substack{k \in \Xi_K \\ \beta = \pm 1}} \eta_{K\beta}(k) \hat{b}^+_{k\beta} \hat{b}_{k\beta}, \qquad (7)$$

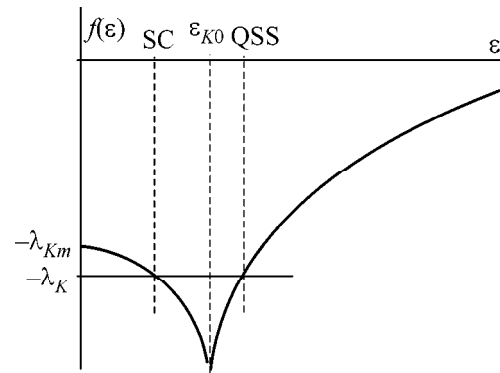**Fig. 2.** Graphical solution of the equation $f(\varepsilon) + \lambda_K = 0$ determining the real part of the pole of scattering amplitude for the relative motion of a hole pair in the case of repulsive interaction of holes [1, 2]. One of the solutions (SC) disappears at $\lambda_K < \lambda_{Km}$ and corresponds to a superconducting condensate of pairs [2], and the second one (QSS) corresponds to quasistationary states of pairs [1].

where

$$\eta_{K\beta} = \left[\varepsilon\left(\frac{K}{2} + \beta k\right) - \mu_K\right] u_{Kk}^2$$

$$- \left(\varepsilon\left(\frac{K}{2} - \beta k\right) - \mu_K\right) v_{Kk}^2 - 2\Delta_{Kk} u_{Kk} v_{Kk}. \tag{8}$$

The nondiagonal part (bilinear in the operators $\hat{b}$) takes the form

$$\hat{H}_K^{(1)} = -\sum_{k \in \Xi_K} \{2\xi_{Kk} u_{Kk} v_{Kk}$$

$$- (u_{Kk}^2 - v_{Kk}^2)\Delta_{Kk}\}(\hat{b}_{k,+1}^+ \hat{b}_{k,-1}^+ + \hat{b}_{k,-1} \hat{b}_{k,+1}). \tag{9}$$

The minimization of E_{K0} with respect to u_{Kk} and v_{Kk} under the additional condition $u_{Kk}^2 + v_{Kk}^2 = 1$ leads to the cancellation of the $\hat{H}_K^{(1)}$ contribution; that is, the braced expression in Eq. (9) identically goes to zero. The equation obtained is fulfilled identically for a solution taken in the form

$$v_{Kk}^2 = \frac{1}{2}\left(1 - \frac{\xi_{Kk}}{\sqrt{\xi_{Kk}^2 + \Delta_{Kk}^2}}\right),$$

$$u_{Kk}^2 = \frac{1}{2}\left(1 + \frac{\xi_{Kk}}{\sqrt{\xi_{Kk}^2 + \Delta_{Kk}^2}}\right). \tag{10}$$

Then, the equation determining Δ_{Kk} takes the form conventional for BCS

$$\Delta_{Kk} = \frac{1}{2}\sum_{k' \in \Xi_K} \frac{U(k - k')\Delta_{Kk'}}{\sqrt{\xi_{Kk'}^2 + \Delta_{Kk'}^2}}. \tag{11}$$

The d-type orbital symmetry of the superconducting gap observed in high-T_c superconducting cuprates follows from Eq. (11) if, as noted in [1], it is considered that Δ_{Kk} as a function of the pair quasimomentum K corresponds to the irreducible representation B_{1g} of the crystal symmetry group.

At $\Delta_{Kk} \longrightarrow 0$, the system can be formally considered [1] as a gas of hole pairs (with a zero binding energy). In this case, the transformation defined by Eq. (3) leads to "correct" zero-order (with respect to the interaction of holes forming a pair) wave functions. Not the entire region Ξ_K but only its part in which $\xi_{Kk} < 0$ contributes to these wave functions. That is, these wave functions are constructed from the states with a positive kinetic energy of the motion of a pair in which the positive heavy component of the effective mass is predominant.

The Lagrange multiplier is determined from the condition

$$2\sum_{k \in \Xi_K} v_{Kk}^2 = 2\sum_{\substack{k \in \Xi_K \\ (\xi_{Kk} < 0)}} 1 \equiv \langle N_K^{(c)}\rangle, \tag{12}$$

where $\langle N_K^{(c)}\rangle$ is the mean number of pairs in the condensate at $\Delta_{Kk} \longrightarrow 0$. The asymmetry in the distribution of states with effective masses of opposite sign within the Ξ_K region can be characterized by the parameter δ_K

$$2\delta_K = \sum_{\substack{k \in \Xi_K \\ (\xi_{Kk} > 0)}} 1 - \sum_{\substack{k \in \Xi_K \\ (\xi_{Kk} < 0)}} 1 = \sum_{k \in \Xi_K} \frac{\xi_{Kk}}{\sqrt{\xi_{Kk}^2 + \Delta_{Kk}^2}}. \tag{13}$$

The parameter in Eq. (13) can, in principle, have any sign. However, one can see from Fig. 1a, where the Ξ_K region typical for high-T_c superconducting compounds with the hole type of doping is shown, that $\delta_K > 0$ at a sufficiently strong anisotropy of the dispersion law of holes, as observed in high-T_c superconducting cuprates [6]. Passing in Eq. (13) from summation to integration over the Ξ_K region and replacing the density of states for the relative motion of pairs within this region by some averaged value g_K [a logarithmic singularity occurring in the density of states $g_K(\xi)$ at $\xi = 0$ only slightly affects the values of integrals within Ξ_K], we find that the Lagrange multiplier takes the form $\mu_{K0} = \delta_K/g_K$ at $\Delta_{Kk} = 0$. As in [2], it is assumed for simplicity that the pair energy (in the absence of interaction) varies within symmetric limits $|\varepsilon(k_+) + \varepsilon(k_-) - 2\varepsilon(K/2)| \leq 2\varepsilon_{K0}$, where ε_{K0} is the maximum kinetic energy of the relative motion of a hole pair within the Ξ_K region (Fig. 1b).

It is evident that a nontrivial solution $\Delta_{Kk} \neq 0$ also exists. In this case, $\Delta_{Kk} \neq 0$ throughout the Ξ_K region, and the part of this region in which states with a negative effective mass along the k_2 axis predominate (that is, bound states can form [1]) leads to a change in the Lagrange parameter. We obtain approximate solutions of Eqs. (11) and (13) by making a radical simplifying assumption that $U(k - k') = U_K = $ const within the Ξ_K region, which is equivalent to the BCS approximation. The consequence of this simplification is that the parameter Δ_{Kk} becomes independent of k, $\Delta_{Kk} \Rightarrow \Delta_K$. This allows the change in the Lagrange parameter $\Delta\mu_K$ upon the formation of a pair condensate to be found from Eq. (13)

$$\Delta\mu_K = \frac{\mu_{K0}}{\varepsilon_{K0}^2 - \mu_{K0}^2} \frac{\Delta_K^2}{2}. \tag{14}$$

With this approximation, the solution of Eq. (11) can be written as

$$\Delta_K = 2\sqrt{\varepsilon_{K0}^2 - \mu_{K0}^2}\exp\left(-\frac{1}{U_K g_K}\right). \qquad (15)$$

Here, it is taken into account that $|\Delta_K| \ll |\varepsilon_{K0} \pm \mu_{K0}|$, which formally corresponds to the weak coupling approximation $U_K g_K \ll 1$.

Rewrite the ground state energy of the system as

$$E_{K0} = -\sum_{k \in \Xi_K} \varepsilon_{Kk}\left(1 - \frac{\xi_{Kk}}{\sqrt{\xi_{Kk}^2 + \Delta_{Kk}^2}}\right) \\ + \frac{1}{2}\sum_{k \in \Xi_K}\frac{\Delta_{Kk}^2}{\sqrt{\xi_{Kk}^2 + \Delta_{Kk}^2}}. \qquad (16)$$

At $\Delta_K \neq 0$, the quasiparticle energy is reckoned from $\mu_K = \varepsilon(K/2) + \mu_{K0} + \Delta\mu_K$, and the ground-state energy [Eq. (16)] of the superconducting phase at $\mu_{K0} \ll \varepsilon_{K0}$ can be written, to an accuracy of $\sim\Delta_K^2$, as

$$E_{K0}^{(s)} \approx E_{K0}^{(n)} - \frac{1}{2}g_K(\Delta_{Km}^2 - \Delta_K^2). \qquad (17)$$

Here, $\Delta_{Km}^2 \approx 4\mu_{K0}\varepsilon_{K0}$, and $E_{K0}^{(n)} = g_K\varepsilon_{K0}^2$ is the ground-state energy of the normal phase in the case when the states of the relative motion of a pair throughout the Ξ_K region are uniformly occupied [in this case, $\langle N_K^{(n)}\rangle = 1$, which corresponds to choosing the coefficients in Eq. (12) in the form $v_{Kk}^2 \equiv v_K^2 = 2\pi^2/S\Xi_K$, where S is the normalization area]. Thus, in the weak coupling approximation and under the condition $\Delta_K < \Delta_{Km}$, it turns out that $E_{K0}^{(s)} < E_{K0}^{(n)}$; i.e., the formation of a condensate provides a gain in the ground-state energy.

The condensation energy can be determined as $E_{K0}^{(c)} = E_{K0}^{(n)} - E_{K0}^{(s)}$ and written in the form of a sum over the Ξ_K region using Eq. (16). A plot of the function $F(x)$ under the corresponding sum (in Δ_K units) against the dimensionless variable $x = \xi_{Kk}/\Delta_K$ is schematically shown in Fig. 3. This figure illustrates how the shift of the chemical potential $\Delta\mu_K$ caused by the formation of a superconducting condensate provides a gain in the condensation energy on account of the kinetic energy of pairs. It is seen in Fig. 3 that a gain corresponding to positive values of $F(x)$ is achieved when the maximum kinetic energy of a pair ε_{K0} [this value determines the limits of integration in Eq. (16)] exceeds a certain value at which the contribution of repulsive interaction (which leads to a loss in the condensation energy) is cancelled (the region in Fig. 3 in which $F(x) < 0$).
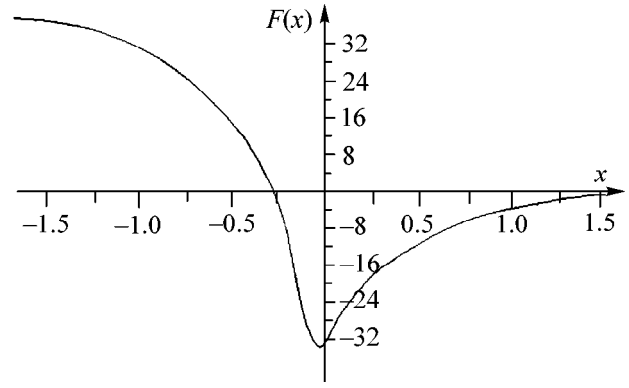
Fig. 3. Condensation energy (in Δ_K units) as a function of the dimensionless excitation energy $x = \xi_{Kk}/\Delta_K$. The plot in the figure corresponds to the shift of the chemical potential of a quasiparticle $\Delta\mu_K/\Delta_K \approx 0.4$.

The phenomenological picture with a gain in the superconducting state energy on account of the kinetic energy was presented in [7].

In the limiting case of weak coupling, Δ_K decreases exponentially at $\lambda_K^{-1} \equiv U_K g_K \longrightarrow 0$, and the pole of the scattering amplitude, which is determined [2] from the condition $f(\varepsilon) + \lambda_K = 0$, shifts toward the point ε_{K0} as shown in Fig. 2. In this case, the decay remains negative and tends to zero at $\varepsilon \longrightarrow \varepsilon_{K0}$. It is evident in Fig. 2 that, with an increase in the coupling constant λ_K^{-1}, the solution of the equation $f(\varepsilon) + \lambda_K = 0$ disappears starting at a certain value λ_{Km}^{-1}. Nevertheless, the parameter Δ_K does not go to zero; moreover, it increases with increasing λ_K^{-1}. However, with increasing Δ_K at a constant ε_{K0}, the positive contribution to the ground-state energy increases, whereas the negative contribution simultaneously decreases, both because of the "contraction" of the limits of integration $\pm\varepsilon_{K0}/\Delta_K$ over the dimensionless variable z. It is seen in Fig. 3 that the contribution of the region with $F(x) > 0$ decreases much faster than the contribution of the region with $F(x) < 0$ (the latter contribution leads to a loss in the condensation energy). As a result, as Δ_K increases, the situation is attained in which $E_{K0}^{(s)} > E_{K0}^{(n)}$. Thus, the disappearance of the instability of the normal state at λ_{Km}^{-1} [2] corresponds to the disappearance of the gain in the energy of the superconducting state over the normal state. As λ_K^{-1} tends to λ_{Km}^{-1}, the superconducting transition temperature T_c must tend to zero. The low temperature T_c and the large value of Δ_K in this region of variation of coupling constant λ_K^{-1} are apparently the reasons for the large value of the ratio $2\Delta_K/T_c$ and the developed quantum fluctua-

tions that are observed in underdoped cuprate superconductors [8].

The existence of a saddle point in the vicinity of the Fermi level was used in [9] to explain the peak in inelastic neutron scattering observed in high-T_c superconductors.

REFERENCES

1. V. I. Belyavskiĭ, V. V. Kapaev, and Yu. V. Kopaev, Zh. Éksp. Teor. Fiz. **118**, 941 (2000) [JETP **91**, 817 (2000)].

2. V. I. Belyavskiĭ and Yu. V. Kopaev, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 734 (2000) [JETP Lett. **72**, 511 (2000)].

3. Z.-X. Shen, W. E. Spicer, D. M. King, *et al.*, Science **267**, 343 (1995).

4. P. Fulde and R. A. Ferrell, Phys. Rev. **135**, A550 (1964).

5. A. I. Larkin and Yu. N. Ovchinnikov, Zh. Éksp. Teor. Fiz. **47**, 1136 (1964) [Sov. Phys. JETP **20**, 762 (1964)].

6. M. Chiao, R. W. Hill, C. Lupien, *et al.*, Phys. Rev. B **62**, 3554 (2000).

7. M. R. Norman, M. Randeria, B. Janko, and J. C. Campuzano, Phys. Rev. B **61**, 14742 (2000).

8. W. Zheng, Phys. Rev. B **62**, 4042 (2000).

9. A. A. Abrikosov, Phys. Rev. B **57**, 8656 (1998).

Translated by A. Bagatur'yants

Purely Antiferromagnetic Oscillatory Mode in a Two-Sublattice Ferromagnetic Phase

E. A. Turov

Institute of Metal Physics, Ural Division, Russian Academy of Sciences, ul. S. Kovalevskoĭ 18, Yekaterinburg, 620219 Russia
e-mail: turov@imp.uran.ru
Received October 27, 2000; in final form, December 8, 2000

It is shown that a purely antiferromagnetic oscillatory mode can be excited by an ac electric field in a two-sublattice single-position ferromagnet. The oscillation eigenfrequency and the corresponding susceptibility tensor are calculated. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 75.25.+z; 75.30.Ds

In considering spin-wave representations for a single-position four-sublattice magnet (four magnetic atoms all occupy crystallographically identical positions in the unit cell), it was shown that, besides the known oscillation types (ferromagnetic in the ferromagnetic phase and quasi-ferromagnetic and quasi- antiferromagnetic in the antiferromagnetic phase), the purely antiferromagnetic modes can also exist in this system [1] (see also [2]). Each of the first two types contains two components of ferromagnetic vector M (precession!) in addition to the dynamic (oscillatory) variables, while only a single component M oscillates in the third type; the remaining variables relate to the antiferromagnetic vectors L_i ($i = 1, 2, 3$). In a purely antiferromagnetic mode, all variables of the spin-wave representation relate to the antiferromagnetic vector, so that these oscillatory modes cannot be excited by a magnetic field $H(t) = H_1 \exp(-i\omega t)$ (there is nothing for them to "clutch at"). One should thus search for the other (other than thermal motion) external perturbations that can excite these modes. Under certain conditions (see below), these may be the ac electric fields and elastic stress.

For brevity, the purely antiferromagnetic excitations will be referred to as antimagnons.

In my opinion, the existence of antimagnons in a purely ferromagnetic phase with the parallely directed magnetization vectors of all sublattices in the ground state is an interesting and, to some extent, exotic fact. It is the purpose of this letter to consider the antimagnons for this particular case.

From the text of articles [1, 2], an impression might occur that the antimagnons (although they were not considered there!) can exist only in the multisublattice magnets (four or more magnetic atoms per unit magnetic cell). However, such is actually not the case. Below, a simple example of a crystal with the twofold position of magnetic atoms (two magnetic sublattices) is presented, in which the purely antiferromagnetic

mode (antimagnon) exists in the ferromagnetic phase. It is reasonable to start investigation of the antimagnons with this very simple example.

Let us consider a twofold 2i position $1(x00)$ and $2(\bar{x}00)$ in the orthorhombic group $Pmmm$ (D_{2h}^1), in which magnetic atoms 1 and 2 are related by the center of symmetry $\bar{1}$ (taken as the origin). There are a total of six positions of this type in the group ($i, j, k, l, m,$ and n) [3], and the subsequent analysis can be reproduced for each of them. The same can also be done for the twofold positions e and f in another orthorhombic group $Pnma$ (D_{2h}^5).

Introducing, instead of the sublattice magnetizations M_1 and M_2, the usual basis ferromagnetic $M = M_1 + M_2$ and antiferromagnetic $M_1 - M_2$ vectors corresponding to the exchange magnetic structures coded $\bar{1}(+)2_x(+)2_y(+)$ and $\bar{1}(-)2_x(+)2_y(-)$, it is straightforward to construct the following table for the variable transformations under the indicated symmetry operations taken as generators of the $Pmmm$ group. (The transformations for the bracketed $2_z = 2_y \cdot 2_x$ element are given for convenience.)

This table contains all symmetry information that is necessary for further consideration. The rows Γ_1–Γ_6 in the upper left part of the table give the one-dimensional irreducible representations of the $Pmmm$ group in terms of the components of vectors M and L. The + and – signs indicate how these components are transformed (either remain unchanged or change their sign) under the corresponding symmetry operation. Each row corresponds to a phase (magnetic structure) specified by a single nonzero component of vector M or L. For example, $M \parallel Z$ for the $\Gamma_3(M_z)$ phase while all remaining components of M or L are zero. At the right, the magnetic point group is indicated for each phase in terms of

the unprimed m_i ($i = x, y, z$) and primed $m_i' = m_i \cdot 1'$ symmetry planes (1' denotes the time inversion operation, $t \longrightarrow -t$). The transformation rules are also given for the components of electric field E and elastic stress tensor t_{ij}.

Using the rules formulated in [1, 2], one can determine the spin-wave representations for each of the six phases. There are two spin-wave representations for the above-mentioned purely ferromagnetic $\Gamma_3(M_z)$ phase of interest: a purely ferromagnetic mode $\Gamma_{12}(M_x, M_y)$ (magnon) and a purely antiferromagnetic mode $\Gamma_{45}(L_x, L_y)$ (antimagnon).[1]

The oscillatory spectrum for a homogeneous system can be found from the Onsager equations

$$\dot{x}_n = \alpha_{nk} \frac{\partial \Phi_2}{\partial x_k} \qquad (1)$$

(summation goes over the repeated indices), where $x_1 \equiv L_x$ and $x_2 \equiv L_y$ are the dynamic variables of the Γ_{45} mode and Φ_2 is the corresponding quadratic part of the thermodynamic potential. The kinetic tensor $\alpha_{nk}(M_z)$ is a function of the ground state with $M_z = M_0$. In the absence of dissipation, only the antisymmetric part of the α_{nk} tensor should be taken into account:

$$\alpha_{nk} \equiv \alpha_{nk}^a(M_0) = -\alpha_{kn}^a(M_0) = \alpha_{kn}^a(-M_0) \cong \gamma M_0. \quad (2)$$

In Eq. (2), the Onsager relation is also used and only the linear term is kept in the expansion of α_{nk} in powers of M_0; γ is a constant for which one can only say that it has the dimensionality of gyromagnetic ratio.

Let us now write the thermodynamic potential (see the table) in the form

$$\Phi_2 = \frac{1}{2}\varepsilon_1 L_x^2 + \frac{1}{2}\varepsilon_2 L_y^2 - \frac{h}{M_0} M_z L_x E_z(t), \qquad (3)$$

which includes the dynamic variables L_x and L_y relevant to the Γ_{45} mode. The ε_1 and ε_2 constants describe both exchange (for which $\varepsilon_1 = \varepsilon_2$) and magnetic anisotropy. The last term in Eq. (3) is a linear magnetoelectric (ME) interaction (the $M_0 = |M_z|$ term in the denominator makes the h constant dimensionless, similar to $\varepsilon_{1, 2}$). One can see from the table that this is the only ME invariant corresponding to the ground ferromagnetic state with $M \parallel Z$. As will be seen below, it is this interaction which can excite antimagnons at the appropriate ω frequency if $E_z(t) \propto \exp(-i\omega t)$.

Equations (1), with allowance made for Eqs. (2) and (3), take the form

$$\dot{L}_x = \gamma M_0 \varepsilon_2 L_y, \quad \dot{L}_y = -\gamma M_0(\varepsilon_1 L_x - hE_z). \qquad (4)$$

The corresponding solution $L_x \propto L_y \propto E_z \propto \exp(-i\omega t)$ defines the antimagnon susceptibility tensor β_{ij} in the relationships

$$L_x = \beta_{xz} E_z, \quad L_y = \beta_{yz} E_z. \qquad (5)$$

Only two components of this tensor are nonzero:

$$\beta_{xz} = \frac{\omega_M^2 \varepsilon_2 h}{\omega_0^2 - \omega^2}, \quad \beta_{yz} = -i\omega \frac{\beta_{xz}}{\omega_M \varepsilon_2}, \quad \omega_M = \gamma M_0. \quad (6)$$

Here, the quantity

$$\omega_0^2 = \omega_M^2 \varepsilon_1 \varepsilon_2 \qquad (7)$$

defines the square of "antimagnon resonance" frequency, whence it follows that the latter has an exchange nature.

Since β_{xz} and β_{yz}, according to Eq. (6), are shifted in phase by $\pi/2$ and, generally, have different magnitudes, the end of the $\mathbf{L}$ vector describes an ellipsis around the vector $\mathbf{M} \parallel \mathbf{Z}$. Near the resonance frequency $\omega \cong \omega_0$, the ratio

$$\frac{L_x}{L_y} = \frac{\beta_{xz}}{\beta_{yz}} = i\frac{\omega_M \varepsilon_2}{\omega} \cong i\sqrt{\frac{\varepsilon_2}{\varepsilon_1}}.$$

This implies that, in a uniaxial crystal considered as a particular case of the orthorhombic crystal with $\varepsilon_1 = \varepsilon_2$, the elliptic polarization becomes circular.

The dissipation, when taken into account, determines heat loss Q associated with the antimagnon excitation by the $E_z(t)$ field. Since the coefficient of $E_z(t)$ in Eq. (3) plays the role of electric induction (displacement), one has

$$Q = \overline{\frac{d\Phi}{dt}} = -h\overline{L_x \frac{dE_z(t)}{dt}}, \qquad (8)$$

where the capping bar denotes time averaging over $t \gg 2\pi/\omega$.

Setting $E_z(t) = E_z \exp(-i\omega t) + E_z^* \exp(i\omega t)$ (and similarly for L_x), one obtains from Eq. (8) with the use of Eqs. (5) and (6)

$$Q = -i\omega h(\beta_{xz} - \beta_{xz}^*)|E_z|^2. \qquad (9)$$

Needless to say, $Q = 0$ if there is no dissipation and, hence, β_{xz} is real, according to Eq. (6). As usual, the damping can be most simply included by substituting in Eqs. (4) $\omega^2 \longrightarrow \omega^2 - i\Gamma\omega$, where Γ is the damping parameter.[2]

---

[1] Recall how the modes are found for the phase under consideration [1, 2]: to a certain mode one attributes the variables of those two rows for which the products of transformations (+1 or –1) yield a row corresponding to the phase under consideration. For the Γ_3 phase, one has: the product of (++−) (M_x) and (+−+) (M_y) is (+−−), which corresponds to M_z, and the product of (−++) (L_x) and (−−−) (L_y) is also (+−−).

[2] One can readily verify that this result can be obtained by adding to the right-hand sides of Eqs. (4) the terms proportional, respectively, to L_x and L_y (as in the case of Bloch equations in paramagnetic resonance [4]).

Table of symmetry operations, representations, phases, and magnetic point groups

Γ_i	M, L	$\bar{1}$	2_x	2_y	(2_z)	Magnetic point group	t_{ij}
Γ_1	M_x	+	+	−	−	$m_x m'_y m'_z$	t_{yz}
Γ_2	M_y	+	−	+	−	$m'_x m_y m'_z$	t_{xz}
Γ_3	M_z	+	−	−	+	$m'_x m'_y m_z$	t_{xy}
Γ_4	L_x	−	+	+	+	$m'_x m'_y m'_z$	
Γ_5	L_y	−	−	−	+	$m_x m_y m'_z$	
Γ_6	L_z	−	−	+	−	$m_x m'_y m_z$	
	E_x	−	+	−	−	$M = M_1 + M_2$	
	E_y	−	−	+	−	$L = M_1 - M_2$	
	E_z	−	−	−	+		

Finally,

$$\beta_{xz} = \beta'_{xz} + i\beta''_{xz}, \qquad (10)$$

where

$$\beta'_{xz} = \frac{(\omega_0^2 - \omega^2)\omega_M^2 \varepsilon_2 h}{(\omega_0^2 - \omega^2)^2 + (\Gamma\omega)^2},$$

$$\beta''_{xz} = \frac{\Gamma\omega\omega_M^2 \varepsilon_2 h}{(\omega_0^2 - \omega^2)^2 + (\Gamma\omega)^2}. \qquad (11)$$

Equation (9) then gives

$$Q = 2\omega h \beta''_{xz} |E_z|^2.$$

This letter, to some extent, may be considered as a tentative prediction of a new type of spin waves (antimagnons) that are excited by an electric field rather than by a magnetic field, as well as some of their properties. The point is that magnets with the magnetic structure considered in this letter are unknown to me (even if it is taken into account that a purely ferromagnetic phase can be prepared artificially by applying a strong magnetic field). For this reason, the system considered is a very simple model that realizes the predicted phenomena. In particular, one can see from it that for this phenomena to exist the magnet must contain magnetic ions related to each other by center of symmetry $\bar{1}$, as is typical of the systems with magnetoelectric interactions. Besides, the interaction of the form $M_i L_j E_k t_{mn}$ can play the active part in such systems (this interaction may be called piezomagnetoelectric).

The table in this letter indicates the existence of $M_z L_y E_y t_{xz}$ and $M_z L_y E_x t_{yz}$ invariants in the thermodynamic potential, as a result of which the antimagnons can be excited in the $\Gamma_3(M_z)$ phase by electric field $E_y(t)$ or $E_x(t)$, provided that a constant elastic stress t_{xz}^0 or t_{yz}^0 is additionally applied. On the other hand, inhomogeneous antimagnons can be excited by the inhomogeneous elastic stress $t_{xz}(t)$ or $t_{yz}(t)$, provided that a constant electric field E_y^0 or E_x^0 is applied.

These and other issues arising in a more complicated (four-sublattice) system will be considered elsewhere in a separate paper.

REFERENCES

1. E. A. Turov and A. V. Kolchanov, in *Magnetism of Transition Crystals and Alloys* (Ural. Otd. Ross. Akad. Nauk, Yekaterinburg, 2000).

2. E. A. Turov, A. V. Kolchanov, V. V. Men'shenin, *et al.*, Usp. Fiz. Nauk **168**, 1303 (1998) [Phys. Usp. **41**, 1191 (1998)].

3. *International Tables for X-ray Crystallograhy* (Kynoch Press, Birmingam, 1983), Vol. 1.

4. C. P. Slichter, *Principles of Magnetic Resonance* (Springer-Verlag, Berlin, 1978; Mir, Moscow, 1981).

Translated by V. Sakun

Induced Jahn–Teller Transition in DyPO$_4$ Caused by a Change in Ground State upon Level Crossing

Z. A. Kazeĭ* and V. V. Snegirev

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia
* e-mail: kazei@plms.phys.msu.su
Received October 10, 2000; in final form, December 13, 2000

The transition caused by a change in the ground state of a rare-earth ion upon level crossing, an analogue of the induced Jahn–Teller transition in rare-earth compounds, is observed. It is shown that a jumpwise change in the Jahn–Teller interactions of α and γ symmetries, caused by a change in the corresponding quadrupole moments in DyPO$_4$ upon level crossing, diminishes the critical field by ~20 kOe and leads to a sharper change in magnetization M and differential magnetic susceptibility dM/dH near the crossover point. © *2001 MAIK "Nauka/Interperiodica"*.

PACS numbers: 71.70.Ej; 75.30.Kz

In recent years, new impressive discoveries in solid state physics in the fields of high-temperature superconductivity and colossal magnetoresistance of compositionally complex perovskites containing Jahn–Teller (JT) Cu^{2+} and Mn^{3+} ions have rekindled interest in the JT systems. Investigations into the spontaneous and induced phase transitions associated with orbital and magnetic ordering, charge-carrier localization, etc., in these compounds have now become the hottest problem of solid state physics. Similar problems, in particular, the spontaneous and magnetic-field-driven structural phase transitions of JT nature (quadrupolar ordering), are currently being studied in rare-earth (RE) paramagnets with the zircon tetragonal structure RXO$_4$ (X = P, As, and V; space group $D_{4h}^{19} = I4_1/amd$), which are ideal objects for the examination of the quadrupolar effects [1]. Reliable experimental data on the interaction parameters of these compounds, in conjunction with modern theoretical approaches, allow not only the qualitative study of the quadrupole interaction effects but also their description on a good quantitative level.

Pronounced quadrupolar effects can be observed in the RE compounds only if the quadrupole interaction constants are large while the electronic structure (spectrum and wave functions of the RE ion) exhibits certain specific features. The RE quadrupole interaction constants are the characteristics of a family of isomorphous compounds and vary, following the Stevens parameter α_J, in a known manner along the RE series. Therefore, the presence or absence of the quadrupolar ordering or, more generally, of the observable quadrupolar effects in RE zircons is dictated by the electronic structure of the RE ion. In the absence of a magnetic field, only those RE zircons show cooperative Jahn–Teller effect which are characterized by a high quadrupole moment

in the low-lying levels (TbVO$_4$, DyVO$_4$, TmVO$_4$, and TbPO$_4$). A strong magnetic field modifies the spectrum and wave functions of the RE ion and may form, under certain conditions, electronic structures with high quadrupole moments, rendering the quadrupolar effects observable.

The electronic structure of RE compounds changes cardinally upon level crossing (crossover). An analysis shows that the level crossing or, more generally, avoided crossing of RE energy levels occurs in strong and ultrastrong magnetic fields for almost all RE vanadates and phosphates having the zircon structure. The crossover was observed and studied in detail for RE zircons HoVO$_4$ [2, 3], YbPO$_4$ [4], and PrVO$_4$ [5] in a magnetic field aligned with the tetragonal axis. In this work, the strengthening of quadrupole interaction caused by a change in the ground state upon level crossing is studied experimentally and theoretically for DyPO$_4$, an example of the ideal Ising-type magnet at low temperatures [6, 7].

The magnetization $M(H)$ of the DyPO$_4$ crystal was studied at low temperatures in a field aligned with the hard axis [100]. Measurements were made by the induction method in pulsed magnetic fields up to 300 kOe (field raise time in the pulse was 10 ms). The digital recording of the signals from the measuring $M(t)$ and field $H(t)$ coils was made at 0.02-ms intervals during the field pulse (about 500 points). The decompensation signal from the measuring coil was recorded under the same conditions in the absence of a sample and subtracted by a program. The programmed experimental data processing consisted in calculating smoothed magnetization function $M(H)$ and its derivative $dM(H)/dH$.

The magnetization along the [100] axis is small in weak fields at low temperatures and increases jump-

wise near the critical field $H_c \sim 140$ kOe, approaching saturation in higher fields (Fig. 1). The critical field H_c can be determined more trustworthily from the curves for differential susceptibility dM/dH (inset in Fig. 1). The magnetization in a strong field is close to a value of $\sim 10\mu_B$, corresponding to the total magnetic moment of the Dy^{3+} ion. The $M(H)$ and $dM(H)/dH$ curves exhibit a small hysteresis at low temperatures (see curves at $T = 17$ K). In our opinion, it is not caused by the measurement error. With a rise in temperature, the $M(H)$ dependence starts to smear and the critical field slightly increases.

The Zeeman effect and the $DyPO_4$ magnetization in a magnetic field were calculated with the Hamiltonian including the crystal-field (CF) Hamiltonian H_{CF}, the Zeeman interaction H_Z, and the bilinear H_B and quadrupole H_{QT} interactions written in terms of the Stevens equivalent operators O_n^m in the quadrupolar and molecular-field approximations (for more detail, see, e.g., [8]):

$$H_{CF} = \alpha_J B_2^0 O_2^0 + \beta_J (B_4^0 O_4^0 + B_4^4 O_4^4)$$
$$+ \gamma_J (B_6^0 O_6^0 + B_6^4 O_6^4), \tag{1}$$

$$H_Z = -g_J \mu_B \mathbf{H} \cdot \mathbf{J},$$
$$H_B = -g_J \mu_B \mathbf{H}_B \cdot \mathbf{J}, \tag{2}$$

$$\mathbf{H}_B = n g_J \mu_B \langle \mathbf{J} \rangle \quad (n = \theta/C), \tag{3}$$

$$H_{QT} = -G^\alpha \langle O_2^0 \rangle O_2^0 - G^\gamma \langle O_2^2 \rangle O_2^2. \tag{4}$$

In these expressions, α_J, β_J, and γ_J are the Stevens parameters; g_J and μ_B are the Lande splitting factor and the Bohr magneton, respectively; and B_n^m are the CF parameters. The total quadrupole constants $G^\mu = G_{ME}^\mu + K^\mu = (B^\mu)^2/C_0^\mu + K^\mu$ ($\mu = \alpha$ and γ) are the sums of contributions from both single-ion magnetoelastic B^μ and pair quadrupole K^μ interactions (C_0^μ is the elastic constant in the absence of interactions). Only the α- and γ-symmetry modes are kept in the H_{QT} Hamiltonian, because only these terms are nonzero in the field $H \parallel [100]$. The crystal field for $DyPO_4$ is assumed to be firmly established from numerous experimental data, including neutron inelastic scattering [9]. The CF parameters were borrowed from [10], where they were determined for the lowest multiplet. The bilinear interactions are not small for $DyPO_4$ and are responsible for the antiferromagnetic ordering of the Dy^{3+} ions at $T_N \sim 3.4$ K [11]. The quadrupole interaction parameters of $DyPO_4$ were determined in [10] for all symmetry modes from the measurements of nonlinear susceptibility and parastriction at low temperatures. The corresponding values $\theta = -1.5$ K, $G^\alpha = 1.5$ mK, and $G^\gamma = 4.4$ mK are used in the following calculations.
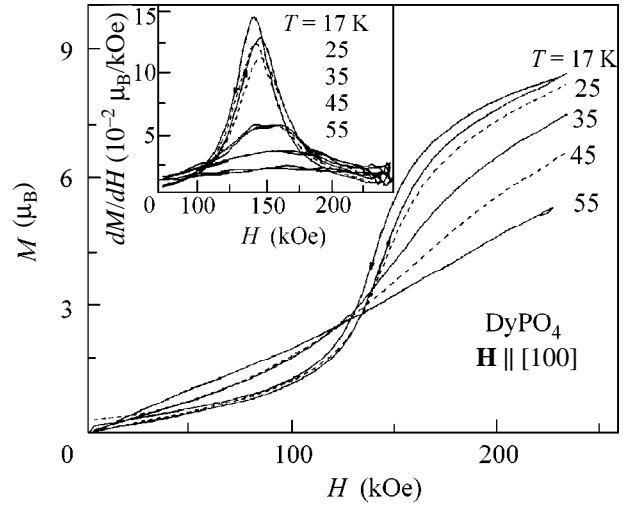
Fig. 1. Magnetization (M) and (inset) differential susceptibility (dM/dH) curves for the $DyPO_4$ crystal in a magnetic field $\mathbf{H} \parallel [100]$ at different temperatures.

In the pulsed fields with a pulse duration of ~ 15 ms, the magnetization proceeds in a near-adiabatic regime and, generally, is accompanied by a considerable magnetocaloric effect. For the sake of comparison with the experiment, the adiabatic and isothermal magnetization curves and the differential susceptibility were calculated together with the magnetocaloric effect. It turned out that the sample heating for the magnetization along the hard [100] axis is small (does not exceed 5 K) at $H < H_c$ and reaches ~ 20 K at $H > H_c$.

The Dy^{3+} spectrum calculated with the above parameters for a magnetic field directed along the [100] axis (Zeeman effect) is presented in Fig. 2. Only six lower levels making the main contribution to the thermodynamic properties at low temperatures are shown in this figure. The lowest Dy^{3+} $^6H_{15/2}$ multiplet is split in the phosphate crystal field in such a way that the "isolated" Kramers doublet with almost the highest possible angular-momentum projection onto the z axis (the smallest onto the x axis) becomes the ground level. The nearest excited doublet is separated by a gap of ~ 100 K and has, on the contrary, the greatest g factor along the x axis ($g_x^{ex} \gg g_z^{ex}$). It is precisely this feature of the Dy^{3+} spectrum which determines its Ising-like character and is favorable to the level crossing for the field orientation along the hard [100] axis.

As expected, the ground doublet with the largest g_z^{gr} component is weakly split in the $\mathbf{H} \parallel [100]$ field, whereas the first excited level is split much more strongly. As a result, the lowest sublevel of the first excited doublet with large $\langle M_x \rangle$ projection approaches in a critical field H_c the ground level with small projec-
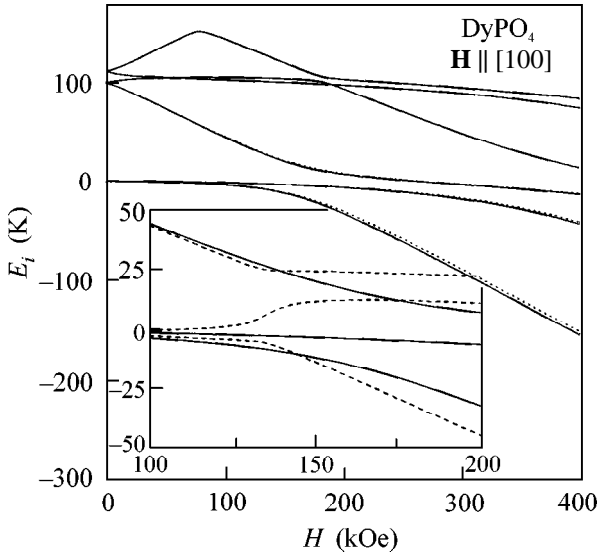
Fig. 2. Energies of the low-lying Dy^{3+} levels in a magnetic field $\mathbf{H} \parallel$ [100] (Zeeman effect) calculated (solid lines) without inclusion of the pair interactions and (dotted lines) with inclusion of the bilinear ($\theta = -1.5$ K) and (inset, dashed lines) quadrupole ($G^{\alpha} = 1.5$ mK and $G^{\gamma} = 4.4$ mK) interactions.

tion of magnetic moment $\langle M_x \rangle$, causing a dramatic increase in the magnetic moment.

A comparison of the spectra calculated with and without taking account of the quadrupole interactions (dashed and solid lines in the inset) indicates that their role appreciably increases above the crossover field. The fact that the spectra coincide in the fields lower than critical means that the Jahn–Teller contribution can be neglected. The calculation with inclusion of only the CF and Zeeman terms yields a value of ~160 kOe for the field H_c corresponding to the level closest approach and to a maximum in the differential susceptibility curve dM/dH (Fig. 3, curve *1*). This is appreciably higher than the experimental value (curve *4*). Note that the demagnetizing field for our sample did not exceed 3 kOe. In view of the fact that the CF parameters are established quite reliably, this discrepancy can be considered as indicative of an additional contribution from the Jahn–Teller interaction.

A sharp increase in this contribution is caused by a change in the quadrupole moments of the Dy^{3+} ion at $H > H_c$. Because of a strong modification of the Dy^{3+} electronic structure, the quadrupole moment $\langle O_2^0 \rangle$ decreases near the crossover point and a large quadrupole moment $\langle O_2^2 \rangle \approx 50$ simultaneously arises. This signifies that the contributions from both terms in the H_{QT} Hamiltonian [Eq. (4)] strongly change. A change of $-G^{\alpha}\Delta Q_0 O_2^0$ [$\Delta Q_0 = \langle O_2^0 \rangle(H) - \langle O_2^0 \rangle(0) < 0$] in the contribution from the α interactions reduces the effective

second-order CF parameter B_2^0 and diminishes the critical field. Similarly, the γ interactions $-G^{\gamma}\langle O_2^2 \rangle O_2^2$, though being negligible in the absence of a magnetic field, noticeably diminish the critical field after the $\langle O_2^2 \rangle$ moment increases jumpwise near the crossover point. Thus, both types of interaction decrease the critical field by approximately 20 kOe.

Note that the contribution from the quadrupole interactions of α symmetry is nonzero even in the absence of a magnetic field and adds $-G^{\alpha}\langle O_2^0 \rangle/\alpha_J$ to the second-order CF parameter B_2^0. Due to a large quadrupole moment $\langle O_2^0 \rangle \approx 100$, this addition amounts to ~10% of the static zero-field parameter B_2^0 at helium temperatures and, with regard to the Stevens parameter α_J, is negative. The experimentally determined low-temperature effective parameter $B_2^0 = \mathscr{B}_2^0 - G^{\alpha}\langle O_2^0 \rangle/\alpha_J$ already includes this addition because it properly describes the low-lying Dy^{3+} levels. The calculations with only the $\mathscr{B}_2^0$ parameter (without inclusion of $G^{\alpha}\langle O_2^0 \rangle/\alpha_J$) would give, e.g., for the energy of the first excited doublet a value greater by ~20 K than its experimental value.

The JT-induced modification of the Dy^{3+} spectrum near the crossover point should alter various thermodynamic characteristics, in particular, the magnetization curves. To estimate the contributions from different interactions, the $M(H)$ and $dM(H)/dH$ curves were calculated with inclusion of the bilinear and α and γ quadrupole interactions (Fig. 3, curves *2* and *3*). As expected, the negative bilinear interactions with $\theta = -1.5$ K slightly increase H_c. By contrast, the quadrupole interactions decrease the critical field and sharpen the jump, while the peak in dM/dH becomes stronger and narrower. The inclusion of the quadrupole interactions with $G^{\alpha} = 1.5$ mK and $G^{\gamma} = 4.4$ mK reduces the critical field down to ~140 kOe and increases twofold the maximum of differential susceptibility, in accordance with the experiment. It should be emphasized that all the calculations were carried out without any fitting parameters and using the values that were determined previously from independent weak-field experiments.

Thus, it is demonstrated in this work that the modification of the electronic structure of a RE ion near the crossover point and a change in the RE ground state can strongly alter the contribution from the quadrupole interactions, leading to the observed quadrupolar effects. At zero magnetic field, the $DyPO_4$ crystal is not a JT compound, contrary to $DyVO_4$, which exhibits spontaneous quadrupolar ordering. This is due to the fact that the CF parameters B_m^n in phosphates differ from those in vanadates, causing a change in the elec-
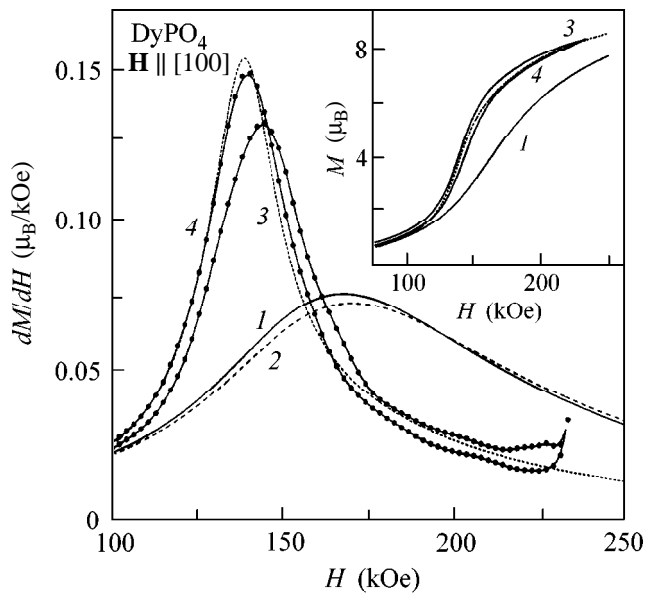
Fig. 3. Experimental (*4*) and calculated adiabatic curves for differential susceptibility *dM/dH* and (inset) magnetization *M(H)* of DyPO$_4$ in a magnetic field **H** ∥ [100] at the initial temperature $T_0 = 17$ K. The calculations are carried out (*1*) without inclusion of the pair interactions and (*2*) with inclusion of the bilinear ($\theta = -1.5$ K) and (*3*) quadrupole ($G^\alpha = 1.5$ mK and $G^\gamma = 4.4$ mK) interactions.

tronic structure of the Dy^{3+} ion in phosphate, where the quadrupole moments of Dy^{3+} are small at $H = 0$. The magnetic field forms the electronic structure in such a way that the quadrupolar effects for the Dy ion near the crossover field H_c in phosphate and at $H = 0$ in vanadate become comparable. A considerable increase (or change) in the quadrupole moments (which are small at zero field) near the crossover point results in an increase in the contribution from the quadrupole terms in the H_{QT} Hamiltonian. Not only the magnetic moment but also the quadrupole moment $\langle O_2^2 \rangle$ increases jump-wise at the critical field H_c, typical of the induced JT transition. However, inasmuch as the magnetic field induces a deformation of the same symmetry as the quadrupole interactions do, the order parameter $\langle O_2^2 \rangle$

of the induced transition is nonzero even in a weak field, while the quadrupole interactions lead to its sharp change in the crossover region. This situation is identical with that observed, e.g., for JT magnetics exhibiting spontaneous transition at $T > T_c$ in a field oriented along the direction of spontaneous crystal deformation [12].

It follows from the calculations that the Zeeman effect for the low-lying levels becomes more complicated upon switching on the quadrupole interactions and strongly depends on the quadrupole interaction constants. In this connection, it is of interest not only to study the thermodynamic characteristics but also to directly examine the Zeeman effect in DyPO$_4$ in the presence of level crossing.

REFERENCES

1. G. A. Gehring and K. A. Gehring, Rep. Prog. Phys. **38**, 1 (1975).
2. T. Goto, A. Tamaki, T. Fujimura, and H. Unoki, J. Phys. Soc. Jpn. **55**, 1613 (1986).
3. P. Morin, J. Rouchy, and Z. Kazei, Phys. Rev. B **51**, 15 103 (1995).
4. Z. A Kazei, N. P. Kolmakova, R. Z. Levitin, *et al.*, Physica B (Amsterdam) **246–247**, 483 (1998).
5. Z. A Kazei, N. P. Kolmakova, V. V. Platonov, *et al.*, Physica B (Amsterdam) **284–288**, 1483 (2000).
6. J. C. Wright, H. W. Moos, J. H. Colwelle, *et al.*, Phys. Rev. B **3**, 843 (1971).
7. I. R. Jahn and S. H. Smith, Phys. Status Solidi B **68**, 531 (1975).
8. P. Morin, J. Rouchy, and D. Schmitt, Phys. Rev. B **37**, 5401 (1988).
9. C.-K. Loong, L. Soderholm, J. Simon Xue, *et al.*, J. Alloys Compd. **207**, 165 (1994).
10. P. Morin and Z. Kazei, J. Phys.: Condens. Matter **11**, 1289 (1999).
11. J. Hoffmann, Phys. Status Solidi B **165**, 517 (1991).
12. P. Morin, J. Rouchy, and Z. Kazei, J. Phys.: Condens. Matter **8**, 7967 (1996).

Translated by V. Sakun

Conductance of a Semiconductor(2DEG)–Superconductor Junction in High Magnetic Field[1]

N. M. Chtchelkatchev

Landau Institute for Theoretical Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia
e-mail: nms@landau.ac.ru

Conductance G of a 2DEG-Superconductor (S) device in a high magnetic field is studied; $G(\nu)$ is calculated. When the cyclotron diameter in 2DEG is larger than the width of the 2DEG-S surface, $G(\nu)$ becomes a non-monotonic function, owing to the Aharonov–Bohm-type interference of quasiparticles at the surface. At certain parameters of the junction, the conductance oscillates with ν. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 85.25.-j; 74.80.Fp

In recent years, the study of hybrid systems consisting of superconductors in contact with normal metals in strong magnetic field has attracted considerable interest [1–5]. Investigation of physical phenomena in S-2DEG devices in high magnetic fields may be helpful in establishing a link between mesoscopic superconductivity and quantum Hall physics. It was found experimentally [1] that zero-bias conductance G of a ballistic S-2DEG-S junction in the Integer Quantum Hall (IQH) regime exhibits quantization upon varying the magnetic field. The quantum of G was not equal to a universal value in this experiment, as, for instance, in IQH or in a quantum point contact [6], but was an oscillating function of field H. Numerical simulations [4, 5] showed that the conductance of a 2DEG-S contact in the IQH regime is a nonmonotonic function of the filling factor ν; the quantization of G is nonuniversal if the 2DEG-S boundary is perfect;[2] in the specific range of magnetic fields, $G(\nu)$ oscillates. A phenomenological theory of conductance oscillations was suggested in [5]. But it is still unclear under which conditions the conductance becomes sensitive to H, why it exhibits oscillations, and how one can analytically describe $G(H)$. The analytical form of $G(\nu)$ is found in this paper. It is shown that the conductance becomes sensitive to H if $2R_c \gtrsim L$, where R_c is a cyclotron radius in 2DEG and L characterizes the length of the 2DEG-S boundary; the nonlinearities in $G(\nu)$ result from Aharonov–Bohm-type interference of quasiparticles at the boundary.

We consider a junction consisting of a superconductor, 2DEG, and a normal conductor segments (Fig. 1). Magnetic field H is applied along the z direction per-

pendicular to the 2DEG plane. It is supposed that the quasiparticle transport is ballistic (mean free path of an electron $l_{tr} \gg L$, where L is the length of the 2DEG-S boundary). Current I is assumed to flow between the normal (N) and superconducting (S) terminals (voltage V is applied across them). The conductance $G(H, L) = I/V$ at $V \longrightarrow 0$ is studied in this paper.

Following [7], we will describe the transport properties of the junction in terms of electron and hole quasiparticle scattering states satisfying the Bogolyubov–de Gennes (BdG) equations. Then, the conductance

$$G = \left.\frac{\partial I}{\partial V}\right|_{V \to 0} = \frac{2e^2}{h}\sum_{l_o, n_i} R_{he, l_o, n_i} = \frac{4e^2}{h}\mathcal{R}, \qquad (1)$$

where R_{he, l_o, n_i} is the probability of Andreev reflection of an electron with energy $E = 0$ (measured from E_f), incident on the superconductor in the channel with quantum number n_i, into a hole moving from the superconductor in the channel l_o.

Before explicit calculation of G, let us discuss, on a qualitative level, how G should depend on H. When H is small ($R_c \ll L$), then $R_{he, l_o, n_i} \simeq R_{he}\delta_{l_o, n_i}$, with R_{he} weakly depending on H. So,

$$\mathcal{R} \simeq R_{he}N, \quad N = Lp_f/2\pi. \qquad (2)$$

If $2R_c \gtrsim L$, the quasiparticles reflected from the superconductor (S) in the normal and Andreev reflection of an electron return to S again because of trajectory bending in a magnetic field. The resulting $G(H)$ dependence is not weak. We will investigate the conductance using the semiclassical approximation $\nu \gg 1$. An electron (hole) quasiparticle in 2DEG can be treated in semiclassics as a beam of rays (in a similar way, light propagation in classical optics is described by beams of rays [8]). Trajectories of the quasiparticle rays can be found

---

[1] This article was submitted by the authors in English.

[2] The S-2DEG boundary is considered "perfect" if the probability of Andreev reflection of an electron or hole quasiparticle with zero energy and momentum directed perpendicular to the surface is close to unity.

from the equations of classical mechanics. If $R_c \gg L$ ($\nu \ll L/\lambda_f$), the edge states in 2DEG do not overlap. Then the quantum numbers n_o, l_i (1) of the incident electron and the reflected hole correspond to the edge states. Reflection of an electron from the superconductor is schematically illustrated in Fig. 1. The electron ray (Fig. 2) reflects from S into the electron and hole rays at y_0. These rays reflect into another hole and electron rays at y_1. So, two hole and two electron rays propagate between y_1 and y_2. Then eight hole rays come from S to N propagating along the same hole path beginning at y_3. One can approximate the probability P of this hole path by the expression

$$
\begin{aligned}
P(y_0, n_i, n_o) &\simeq \left| r_{he} r_{ee} r_{ee} r_{ee} e^{3iS_e - i3\pi/2} \right. \\
&\left. + r_{hh} r_{he} r_{ee} r_{ee} e^{iS_h + 2iS_e - i\pi/2} + \dots \right|^2,
\end{aligned}
\tag{3}
$$

where r_{ba} is the amplitude of reflection of a quasiparticle a into a quasiparticle b from the superconductor. $S_{e(h)}$ is the quasiclassical action of an electron (hole), taken along the part of the trajectory connecting adjacent points of reflection. Then, $R_{eh, l_o, n_i} \simeq \langle P(y_0) \rangle \delta_{n_i, l_o}$, where the average is taken over $0 < y_0 < d(n_i)$, with d being the length of quasiparticle "jump" along the edge of 2DEG (see Fig. 2). Equation (3) includes the interference terms depending on $S_e - S_h$. Since $S_e - S_h = 2\pi(\nu - 1/2)$, one can expect that the conductance is a nonlinear function of ν, due to the interference terms. The nature of this nonlinearity, indeed, resembles the Aharonov–Bohm effect, as it was assumed in [5], for which the conductance oscillates with H because the vector potential changes the phases of electrons moving from a source along different paths. It will be seen below that, under certain conditions, the S-2DEG conductance oscillates with ν.

The above semiclassical estimates assume that spin degeneracy takes place; T, $eV = 0$; and that diffraction is small: the difference in the hole and electron momenta at $E > 0$ is neglected. This approximation is valid if $\max\{T, |eV|, g\mu_B H\}/\mu \ll \lambda_F/L$, where λ_F is the Fermi wavelength in 2DEG. Below, these conditions are also assumed to be satisfied in the conductance calculation.

The conductance of the 2DEG-S structure will be calculated as a semiclassical asymptotic limit ($\nu \gg 1$) of Eq. (1). If the S-2DEG surface is flat, then the approach described in [7, 9] gives an idea how one can express $R_{eh, ln}$ via the semiclassical asymptotic form of the Green's functions of BdG equations. The relevant calculations confirm that the above naive estimates of G actually lead to the semiclassical asymptotic expression for the conductance:

$$
\mathcal{R} = \sum_{n_i} \int_0^{d(n_i)} \left\{ \rho(n_i, y_0) \left| \sum_a t_a \exp\left(iS_a - i\frac{\pi}{2}\mu_a \right) \right|^2 \right\} dy_0, \tag{4}
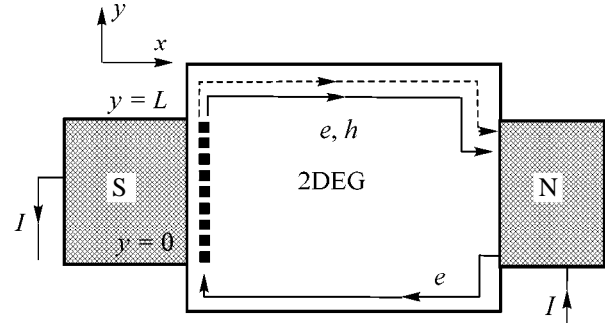$$

Fig. 1. The device under consideration consists of a superconductor, 2DEG, and a normal conductor. An electron injected from the normal conductor in the IQH regime passes through an edge state to the superconductor and reflects into a hole and an electron, which then return to the normal contact through the other edge states.
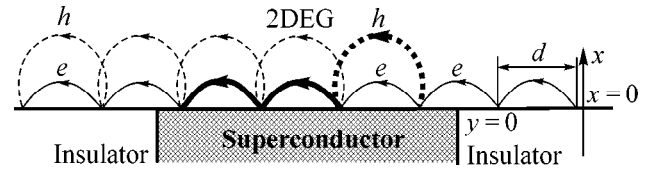

Fig. 2. A quasiparticle in semiclassical approximation can be treated as a beam of rays propagating along classical trajectories. The figure illustrates reflection of rays from the superconductor. Dashed curves correspond to the hole rays, and solid curves, to the electron rays.

where n_i is the index of the edge state of an electron incident on the superconductor; $d(n_i)$ is the quasiparticle jump length; and t_a is the probability amplitude for the classical quasiparticle trajectory from $y = y_0$ to $y = y_{n_0}$, the coordinate of the last reflection from the superconductor. The amplitude t_a is a product of the Andreev and the normal reflection amplitudes; S_a is the action taken from y_0, the coordinate of the first quasiparticle reflection, to y_{n_o}, the coordinate of the last reflection (see Fig. 2); and μ_a is Maslov index of the trajectory. For example, $t_a = r_{he} r_{ee} r_{eh} r_{he}$ and $S_a = S_h + 2S_e$ for the trajectory drawn by the bold line in Fig. 2. Summation over a means the sum over all paths connecting y_0 with y_{n_o} at the 2DEG-S boundary. The weight function $\rho(n_i, y_0)$, where $\int_0^d \rho(n_i, y_0) dy_0 = 1$, generally depends on the shape of the 2DEG-S contact. If 2DEG spreads over the region $x > 0$, $y > 2R_c + L$, $y < -2R_c$, as in Fig. 2, then $\rho = 1/d$. Equation (4) is the key result of this work.

The sum over trajectories in Eq. (4) can be converted into the analytical expression

$$
G = \frac{4e^2}{h} \sum_{n_i} \sum_s P_s \frac{R_{eh} \sin^2(s \arccos(\sqrt{R_{ee}} \cos(\Omega)))}{1 - R_{ee} \cos^2(\Omega)}, \tag{5}
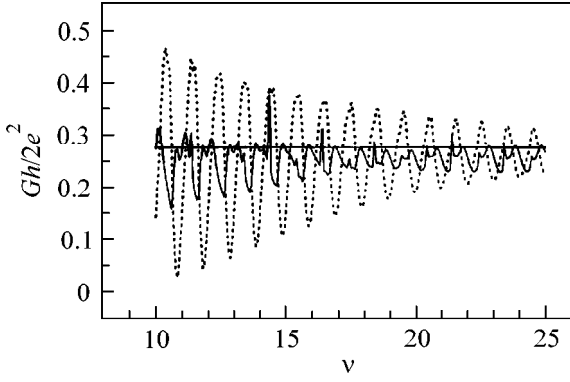$$

Fig. 3. The solid curve represents the conductance for $\lambda L = 3$, and the dots are for $\lambda L = 1$. The line parallel to the X axis is the conductance at $H = 0$.

where $\Omega = \pi\nu + \theta - 2\lambda p_\perp$; $\theta = arg(r_{ee})$ is the phase of the electron–electron reflection amplitude from the superconductor, $R_{ee} = |r_{ee}|^2$; $p_\perp = p_\perp(n_i)$ is the perpendicular momentum component of a quasiparticle at the instant it reflects from the superconductor; λ is the penetration depth of the superconductor; and $\nu = E_f/(\hbar w_c) - 1/2$. Function P_s is the probability that s reflections happened from the surface of the superconductor. If $\rho = 1/d$, this function can be expressed through the maximum number of jumps $g = [L/d]$ across the S-2DEG surface, where [...] denotes the integral part:

$$P_s = \begin{cases} \dfrac{L - gd}{d} & \text{if} \quad s = g + 1 \\[2mm] 1 - \dfrac{L - gd}{d} & \text{if} \quad s = g \\[2mm] 0 & \text{otherwise.} \end{cases} \qquad (6)$$

Equations (4), (5) are the key result of this work; they show how the conductance depends on magnetic field and parameters of the contact. For small magnetic fields, Eq. (5) reduces to Eq. (2). It follows from Eq. (5) that at $2R_c \gtrsim L$ the conductance becomes sensitive to H. Some limiting cases of Eq. (5) will be considered below for this regime. The Aharonov–Bohm-type conductance oscillations are the most interesting property of $G(\nu)$. It follows from Eq. (5) that the oscillations are visible when $\lambda/L \ll R_c^2/L^2$ and $R_{eh} \lesssim 1/2$. A typical contact, for which the S terminal is made from type-I superconductor and 2DEG is formed from GaAs, has $\lambda \sim \lambda_F \sim 10^{-6}$ cm and $L \sim 10^{-3}$ cm. It follows from the conditions given above that, if $R_c \sim L$, the oscillations can be seen in the contact. (These oscillations were numerically investigated in [4, 5]. It was checked that the theory presented in this paper is consistent with the numerical calculations.) When $R_{ee} \ll 1$, the conductance shows step features. It is of interest to investigate

the regime $R_{eh} \ll 1$, $L/R_c \gg 1$. For this regime, the functional dependence $G(\nu)$ resembles the light intensity distribution $I(\delta)$ that is observed in optics with a Lummer–Gerike interferometer [8], if one considers s as the number of beams in the interferometer and $\delta = 2\Omega$ as the phase difference between successive beams. The probability R_{he} will correspond to the transmission probability through the mirror of the interferometer. Examination of the quasiparticle trajectories making the main contribution to the conductance in Eq. (4) shows that they are similar to the trajectories of light beams in the interferometer.

Figure 3 illustrates how the conductance depends on the magnetic field according to Eq. (5). One curve corresponds to $\lambda p_f = 1$ (dots), and the other to $\lambda p_f = 3$ (solid line). The solid line parallel to the X axis is for conductance (2). If $2R_c \lesssim L$, the conductance oscillates because of the quasiparticle interference, but when ν becomes larger ($2R_c \lesssim L$), the interference phenomena become less probable and the conductance approaches (2). The graph also demonstrates that an increase in the ratio λ/L leads to smearing of the oscillations. The following contact parameters were used: $Lp_f = 80$ and $Z = 5$. Parameter Z [10] characterizes the normal scattering from the superconductor caused by Schottky barriers, differences in the effective masses in 2DEG and S, and so on. Amplitudes r_{ee} and r_{eh} of the normal and the Andreev reflections from the superconductor were calculated for zero magnetic field by matching the quasiparticle wave functions in 2DEG with the wave functions in S. This procedure is true if $(\lambda/R_c)^2 \ll 1$, where R_c is the cyclotron radius in the superconductor. Magnetic fields used in the plot satisfy this condition. The ratio Δ_0/E_f was equal to 0.02, and the gap $\Delta(x) = \Delta_0$ in the superconductor and is zero in 2DEG.

It seems important to determine how a disorder at the S-2DEG surface affects $G(\nu)$. The disorder can be caused by surface roughness, impurities, and so on. It may induce fluctuations of S_a and t_a in Eq. (4), thus breaking the quasiparticle interference. In this work, the disorder is characterized by the mean free path l_{tr} of electron elastic scattering from impurities and by the root-mean-square deviation $\delta n = \sqrt{\langle (\mathbf{n} - \mathbf{n}_0)^2 \rangle}$ of the normal unit vector to the S-2DEG boundary from the $\mathbf{n}_0$ direction of the X axis, where $\langle (\mathbf{n} - \mathbf{n}_0)^2 \rangle = \int_0^L dy (\mathbf{n}(y) - \mathbf{n}_0)^2/L$. Then, one can deduce that Eq. (5) is correct if $R_c \ll l_{tr}$ and $\delta n \ll \min\{l_H^2/L\lambda, 1/p_f L\}$.

If the superconductor surface is diffusive, i.e., $\delta n \gtrsim \max\{1/\lambda p_f, 1/\nu\}$ or $l_{tr} \lesssim R_c$, then there is no interference between different paths in Eq. (4). One can estimate the

reflection probability in this regime by neglecting the interference terms in Eq. (5). Then,

$$\mathcal{R} \simeq \sum_{n_i} \int_0^{d(n_i)} dy_0 \sum_a |t_a|^2 / d(n_i). \qquad (7)$$

The number of quasiparticle reflections from the surface of the superconductor is estimated as $s_0 = [L/2R_c]$; then,

$$G = \frac{4e^2}{h} \mathcal{R} \simeq [\nu] \frac{4e^2}{h} \langle R_{eh} \rangle$$
$$\times (1 - 2\langle R_{eh} \rangle)^{1 - s_0/2} U_{s_0}(\langle R_{ee} \rangle)/\sqrt{1 - 2\langle R_{eh} \rangle}, \qquad (8)$$

where $U_s(x) = \sin(s \arccos(x))/\sin(\arccos(x))$ is a Chebyshev polynomial of the second kind [11]. As $s_0 \longrightarrow \infty$, conductance (8) tends to $[\nu]2e^2/h$.

One can assume that the deviations from the ideal conductance quantization (with the universal step $2e^2/h$) observed in experiment [1] are due to the above-mentioned interference of quasiparticles. The S-2DEG boundary of the device used in [1] was hardly flat, so that the contributions from harmonics with large s [see Eq. (5)] should be suppressed. Then, it is reasonable to approximate the conductance G by $(2e^2/h)f(\nu)(1 + a\cos(\pi\nu + \varphi_0))$, where $a \ll 1$ characterizes the disorder at the surface, φ_0 is a phase shift, and $f(\nu) \sim [\nu]$ describes the shape of the conductance "quantum." This formula was found to be a good fit to the experimental data.

REFERENCES

1. H. Takayanagi and T. Akazaki, Physica B (Amsterdam) **249–251**, 462 (1998).

2. T. D. Moore and D. A. Williams, Phys. Rev. B **59**, 7308 (1999).

3. H. Hoppe, U. Zülicke, and G. Schön, Phys. Rev. Lett. **84**, 1804 (2000).

4. Y. Takagaki, Phys. Rev. B **57**, 4009 (1998).

5. Y. Asano, Phys. Rev. B **61**, 1732 (2000); Y. Asano and T. Yuito, Phys. Rev. B **62**, 7477 (2000).

6. L. I. Glazman, G. B. Lesovik, D. E. Khmelnitskii, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **48**, 218 (1988) [JETP Lett. **48**, 238 (1988)].

7. C. J. Lambert, J. Phys.: Condens. Matter **3**, 6579 (1991); Y. Takane and H. Ebisawa, J. Phys. Soc. Jpn. **61**, 1685 (1992).

8. M. Born and E. Wolf, in *Principles of Optics* (Pergamon, Oxford, 1986), p. 341.

9. H. Baranger, D. DiVincenzo, R. Jalabert, *et al.*, Phys. Rev. B **44**, 10637 (1991); K. Richter, *Semiclassical Theory of Mesoscopic Quantum Systems* (Springer-Verlag, Berlin, 2000), Springer Tracts Mod. Phys. **161**, 63 (2000).

10. G. E. Blonder, M. Tinkham, and T. M. Klapwijk, Phys. Rev. B **25**, 4515 (1982).

11. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Nauka, Moscow, 1971; Academic, New York, 1980).

Fluctuation Model of Current-Driven Magnon Excitation

M. V. Tsoĭ[1,2] and V. S. Tsoĭ[1,3]

[1] *Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia*

[2] *Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824-1116, USA*

[3] *e-mail: tsoi@issp.ac.ru*

A model is suggested that explains the stationary high-frequency magnetization oscillations observed previously by M. Tsoĭ *et al.* when passing dc current through a silver tip mounted on a magnetic Co/Cu multilayer. At the interface between nonmagnetic and ferromagnetic (N/F) metals, Aronov's gap Δ_A (difference in the electrochemical potentials of electrons with opposite spins) arises upon the passage of electric current, thereby energetically promoting magnon creation. Electrons flowing from the nonmagnetic to the ferromagnetic metal become spin-polarized. In a magnetic field, magnetization fluctuations in a mesoscopic ferromagnet give rise to magnetization precession around the magnetic field. The precession is damped due to the viscous losses. In the presence of spin-polarized electron flow, the fluctuations also produce a current-induced torque that compensates for the dissipative torque, leading to stationary high-frequency magnetization oscillations. © *2001 MAIK "Nauka/Interperiodica".*

PACS numbers: 75.30.Ds; 72.25.Ba

High-frequency magnetization oscillations (HFMOs, magnons) induced by electric current flowing perpendicular to a magnetic multilayer were predicted by Slonczewski and Berger [1, 2] and observed in [3, 4]. In [3, 4], HFMOs were excited by passing dc current through a microcontact formed by a silver tip mounted on a Co/Cu multilayer composed of 20–50 pairs of ~2-nm-thick layers exposed to a strong magnetic field **H**, which was higher than the Co spontaneous magnetization field and aligned with the normal to the multilayer surface (**H** ∥ **N**). The HFMOs were detected by measuring electrical resistance of the microcontact. The excitation of HFMOs appeared as a stepwise increase in the microcontact resistance and was observed only for one current direction. The HFMOs were also observed, e.g., in [5–7] for multilayers with two ferromagnetic layers. In [5, 7], as in [3], the resistance was measured. The microcontact used in [5] was fabricated by electron-beam lithography and mounted on a Cu/Co film structure exposed to a strong magnetic field **H** ∥ **N**. The observation of the HFMOs in a multilayer with one cobalt layer in zero external magnetic field was also reported in [5]. The multilayer system used in [7] consisted of alternating copper and cobalt disks ~100 nm in diameter, and the HFMOs were observed in a strong magnetic field both parallel and perpendicular to the multilayer surface. In [6], HFMOs were excited by ac current flowing through a tip into a Fe/Cr film structure in a magnetic field **H** ⊥ **N**, and the oscillations were detected using Brillouin light scattering.

In the theoretical works on HFMOs [1, 2], two different models were applied to a five-layer structure of alternating ferromagnetic and nonmagnetic metal layers, namely, three layers of a normal metal and two ferromagnetic layers crossed by a flow of conduction electrons. The conductor cross sections in the layer plane were assumed to be mesoscopic, the external field was taken to be zero, and the ferromagnet magnetizations were noncollinear. The HFMOs result from the magnetization transfer to the ferromagnetic layer by the spin-polarized electron flow carrying (spin) angular-momentum flux. For the magnetization to transfer between two ferromagnets, the following conditions must be fulfilled: (1) the first ferromagnet must polarize spins of the nonpolarized electron flow from the nonmagnetic metal and make the outflowing electrons spin-polarized; (2) the thickness of the normal metal spacer between the ferromagnets is smaller than the spin-relaxation length Λ_{sN} in it, so that the electron flow incident on the second ferromagnet is spin-polarized and the angular-momentum flux into the second ferromagnet is nonzero; (3) the ferromagnet magnetizations must be noncollinear, otherwise no magnetization transfer occurs; and (4) the angular-momentum flux must be absorbed by the second ferromagnet. For identical ferromagnets, the magnetization transfer results in a propeller-like rotation of the ferromagnet magnetizations in their common plane [1, 2]. In [8], the ballistic model [1] was used to develop, under the same basic assumptions, a macroscopic theory of HFMOs excited by passing dc current through a microcontact mounted on a multilayer film containing two ferromagnetic lay-

ers and exposed to a strong magnetic field perpendicular to the film surface.

In real experiments [3–5, 7], the basic conditions for applicability of the theory of current-induced HFMO excitation [1, 2, 8] are broken. In particular, (a) all magnetizations are collinear, and, hence, condition (3) is broken, and (b) the layer thicknesses are much smaller than the spin-relaxation length Λ_{sF} in the ferromagnet, and, hence, conditions (1) and (4) are also broken (about spin relaxation, see, e.g., [9]).

In this work, the experimental situations realized in [3–5, 7] are critically analyzed and a model is suggested that explains the stationary current-induced HFMO excitation in the thin film and disk magnetic multilayers and the HFMO excitation in a single ferromagnetic layer. The model explains the HFMO excitation only for one (and not for the opposite) current direction and the presence of a current threshold for the excitation.

At finite temperature, spin waves are always present in a ferromagnet. The spin-wave amplification (generation) by a flow of charged particles was predicted by Akhiezer, Bar'yakhtar, and Peletminskiĭ as early as 1963 (see [10]). The problem of current-driven magnon excitation amounts to the problem of macroscopic spin-system "buildup." Energetic estimates suggest (see Fig. 2 in [3]) that almost every electron injected into the ferromagnet creates a magnon. Under these conditions, it seems natural to describe the current-driven magnon generation by using the Landau–Lifshitz equation [11].

The key conditions for HFMO excitation in ferromagnet F are as follows: (I) energetically, magnons can be created in F; (II) the conduction electrons incident on F are spin-polarized; (III) the magnetization vector **M** in ferromagnet F and the polarization vector **S** of electron beam are noncollinear; and (IV) the angular-momentum flux dissipates in ferromagnet F.

For condition I to be met, high current density I is necessary. The current assists in creating electronic states with opposite spins. Due to the transitions between these states, the ferromagnet magnetization changes and the energy necessary for creating a magnon is released. Two mechanisms are operative (Fig. 1): (1) the formation of Aronov's gap $\Delta_A(I)$ at the N/F interface [12] (Fig. 1a) and (2) the generation of nonequilibrium excitations with energy $\Delta_\varepsilon(I)$ by electric current [2] (Figs. 1b, 1c).

The reasons for the appearance of a spin-polarized electron beam entering the nonmagnetic metal from the ferromagnet are apparent [12]. The polarization arises as a result of the transformation of the electronic spectrum at the N/F interface under the condition that the electrochemical potential, as well as the electron flow for each of the spin polarizations, is continuous [15]. The problem can exactly be solved for the diffusion conduction mechanism [12]. For the same reasons and under the same conditions, the spin polarization of an electron beam entering the ferromagnet from the non-

magnetic metal (Fig. 1a) is no less apparent. Hence, the electrons flowing into F from the nonmagnetic metal are spin-polarized, and condition II is fulfilled. The magnetization of the nonmagnetic metal near the microcontact is determined by the magnetization of the adjacent region of size on the order of spin-relaxation length, which is much larger than the microcontact size a. The spin polarization **S** of the electron flow is collinear with the magnetization **M** of the contacting F. We assume that the electronic spectrum is transformed and Aronov's gap $\Delta_A(I)$ appears. Such a transformation seems to be natural for a magnetic multilayer with ferromagnetically ordered layers [3, 4]. For an arbitrary experimental situation, the rigid conditions accepted in theory [12] fall far short of being necessarily fulfilled, so that an additional argumentation is needed in each case.

Condition III is fulfilled fluctuatively. The equilibrium magnetization **M*** in the near-contact region of the ferromagnet is collinear with the electron-beam polarization. The nonzero values of the vector product [**M*** × **S**] are due to the fluctuations of magnetization **M***.

The electron density is ~10^{23} el/cm^3. For an experimental current density of ~10^8 A/cm^2, the injected magnetization constitutes ~10^{-4} of the ferromagnet magnetization; i.e., it is only a weak perturbation. The condition for current-driven excitation of F is dictated by the angular-momentum conservation law and amounts to the requirement that the total angular-momentum flux into F be nonzero. One can suppose that the spin flux transferred by the electron flow is zero if the electron spin state does not change during the electron passage through F. For the classical passage of electrons through F, the time scale for a change in the electron spin state in F ("change" in the magnetic moment of an electron) is determined by the Larmor period. This change is small, provided that the time of electron passage through F is much shorter than the Larmor period. For a 2-nm-thick ferromagnet and a Fermi velocity of ~10^8 cm/s, the time of electron ballistic flight is a factor of ~10^4 shorter than the Larmor period (~10^{-11} s [3]). However, it should be taken into account that, first, the desired change in **M** is small and, second, the absorption of angular-momentum flux in a ferromagnet with thickness $\ll \Lambda_{sF}$ may increase due to several mechanisms. Among these are (1) a dramatic decrease in l upon increasing electron-excitation energy (see, e.g., [16]); (2) capture of an electron incident on F into the surface state [17, 18]; and (3) spin-flip scattering by the local magnetic centers at the N/F interface and by the nonmagnetic defects both at the N/F interface and in F. It is assumed that these mechanisms of angular-momentum flux dissipation in a ferromagnet provide the required weak perturbation of the ferromagnetic state; i.e., condition IV is then fulfilled.

The excitation of the HFMOs in a film magnetic multilayer by passing electric current through a microcontact made from a nonmagnetic metal and mounted
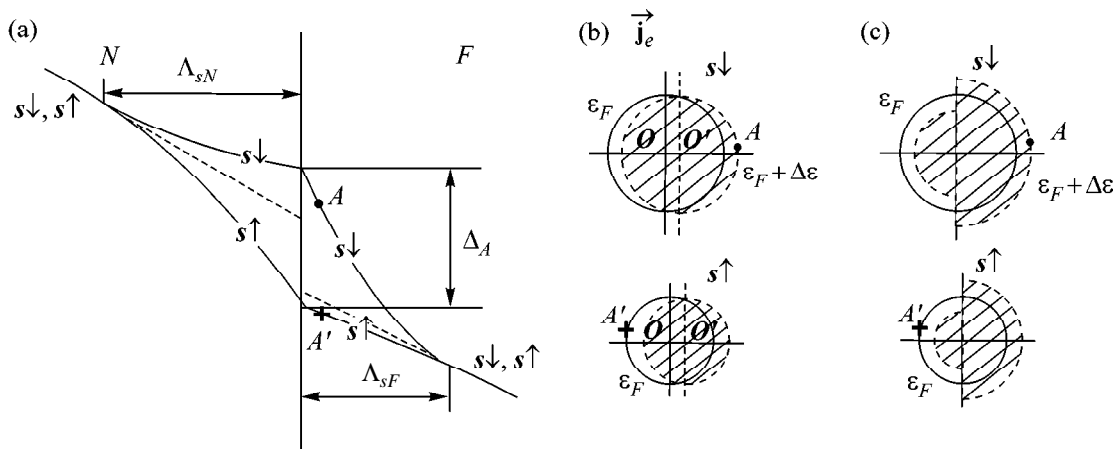
Fig. 1. Excitation of an electron system by the electric current in the (a, b) diffusion and (c) ballistic conduction regimes; $\mathbf{j}_e$, $\mathbf{s}{\uparrow}$, and $\mathbf{s}{\downarrow}$ indicate the electron drift direction and the electron spin polarizations, respectively. Electron transition from the excited state A (circle) to the empty state A' (+) is accompanied by a change in the electron spin polarization and by the energy release. (a) Variation of the electrochemical potential along the normal to the N/F interface (N is nonmagnetic metal, and F is ferromagnet). Due to the spin asymmetry of the heterocontact materials, Δ_A (difference in electrochemical potentials of electrons with opposite spins) arises at the N/F interface [12]. The ferromagnet magnetization should decrease upon the HFMO excitation. In the case that the conductance and the magnetization are determined by the electrons with different spin polarizations (Co), the HFMOs are excited near the surface through which electrons flow into the ferromagnet ($A \longrightarrow A'$ transition). As the current at the N/F interface changes its direction, the electrochemical potential for $\mathbf{s}{\uparrow}$ becomes higher than for $\mathbf{s}{\downarrow}$. If both conductance and magnetization of the ferromagnet are determined by electrons with the same spin polarization, the HFMO excitation occurs near the surface through which electrons outflow from the ferromagnet. The current-induced electron excitation in momentum space is illustrated in Figs. 1b and 1c. The Fermi surface in the equilibrium state is shown by the solid line and the filled region in the excited state is cross-hatched; ε_F is the Fermi energy and $\Delta\varepsilon$ is the excitation energy. (b) The OO' displacement is equal to $e \times \mathbf{E} \times \tau$ (e is the electron charge, $\mathbf{E}$ is the electric field, and τ is the relaxation time), and the current density is $|\mathbf{j}_e| = \rho \times v_D \times \tau$ (ρ and v_D are the electron density and the electron drift velocity, respectively). (c) Ballistic regime of electron injection through the microcontact [13, 14]. The maximum excitation energy is $\sim eV/2$, where V is the microcontact voltage.

on a magnetic multilayer proceeds in the following way (Fig. 2a). A local area of the magnetic layer under the microcontact with characteristic size on the order of a and high current density (magnon disk) plays the role of excited ferromagnet F [3, 8]. This conclusion is consistent with the fact that the microcontact resistance (which is measured to gain information about the multilayer excitation) is determined by a near-contact region with characteristic size $\sim a$ [13, 14]. For simplicity, we assume that the magnon disk contains only a single domain and the current density flowing through it is uniform. The electron beam injected from the nonmagnetic tip into F is spin-polarized, and $[\mathbf{M} \times \mathbf{S}] = 0$. In the presence of fluctuations, the magnetization $\mathbf{M}^*$ of the magnon disk is noncollinear with $\mathbf{S}$; i.e., $[\mathbf{M}^* \times \mathbf{S}] \neq 0$. The magnetic dynamics is determined from the Landau–Lifshitz equation [11]. Let us introduce unit vectors $\mathbf{s}_F$ and $\mathbf{s}_e$ for $\mathbf{M}^*$ and $\mathbf{S}$, respectively. In a magnetic field, the magnetic moment $\mathbf{M}^*$ precesses around the field. Due to the dissipation proportional to the precession velocity, i.e., in the presence of a viscous torque acting on $\mathbf{M}^*$, the precession decays. Within the phenomenological approach, the viscous torque for $|\mathbf{M}^*| =$ const is usually written as a vector product [10, 11, 19]

$$-(\alpha/\gamma)[\mathbf{M}^* \times d\mathbf{M}^*/dt]/[\mathbf{M}^*], \qquad (1)$$

where α is the Hilbert parameter, γ is the electron gyromagnetic ratio, and t is time. The current-induced magnetization flow into the ferromagnet is $(I_e/e)g(\theta)[\mathbf{s}_F \times [\mathbf{s}_F \times \mathbf{s}_e]]$, where I_e is the electric current density and $g(\theta)$ is a positive constant depending on the structure parameters and the angle between $\mathbf{M}^*$ and $\mathbf{S}$ [1, 2]. Assuming that only the conduction electrons are responsible for the magnetization $\mathbf{M}^*$, one can reduce the Landau–Lifshitz equation to the form {see Eq. (15) in [1]}

$$d\mathbf{s}_F/dt = \gamma[\mathbf{s}_F \times \mathbf{B}_{\text{eff}}] - \alpha[\mathbf{s}_F \times d\mathbf{s}_F/dt]$$
$$+ g(\theta)(N/N_0)[\mathbf{s}_F \times [\mathbf{s}_F \times \mathbf{s}_e]], \qquad (2)$$

where $\mathbf{B}_{\text{eff}}$ is the effective magnetic field, N is the number of polarized electrons injected into F in unit time, and N_0 is the difference in the numbers of electrons with different spin polarizations. This difference defines the ferromagnet magnetization. We assume that the electron distribution is properly corrected for electroneutrality. If the condition

$$\alpha d\mathbf{s}_F/dt - g(\theta)(N/N_0)[\mathbf{s}_F \times \mathbf{s}_e] = 0 \qquad (3)$$

is fulfilled, Eq. (2) has a stationary solution in the form of HFMOs: $\mathbf{M}^*(\mathbf{r}, t) = \mathbf{M}_0 + \mathbf{m}\exp[i(\mathbf{k}_m\mathbf{r} - \omega_m t)]$, where $\mathbf{m}$, $\mathbf{k}_m$, and ω_m are, respectively, the amplitude, the wave vector, and the frequency of a magnon; and $\mathbf{r}$ is the coordinate.
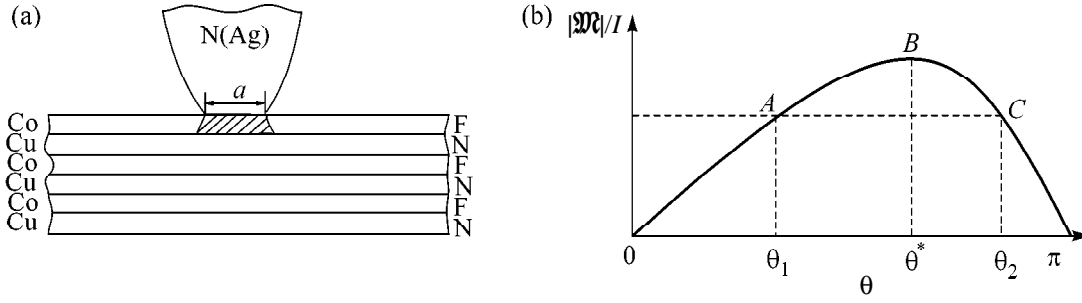
Fig. 2. (a) Scheme of the microcontact HFMO excitation in a magnetic multilayer [3]. $N(Ag)$ is a silver tip, and a is the microcontact diameter. The ferromagnet region where HFMOs are excited (magnon disk) is cross-hatched [3, 8]. The interface area between the magnon disk and the remaining ferromagnet, where the exchange interaction impedes excitation of the magnon disk, increases upon replacing the multilayer by a bulk ferromagnet. (b) Current-induced effective torque $|\mathfrak{M}|$ normalized to current density I (the scale of ordinate axis is arbitrary) vs. angle θ between the ferromagnet magnetization and the electron-beam spin polarization [1]. The point A corresponds to unstable equilibrium: if $\delta\theta$ is added to θ_1, the corresponding increment in $|\mathfrak{M}|$ tends to increase $\delta\theta$. The point C corresponds to stable equilibrium: if $\delta\theta$ is added to θ_2, the corresponding increment in $|\mathfrak{M}|$ tends to decrease $\delta\theta$. The point at which $|\mathfrak{M}|$ is maximum (B, $\theta = \theta^*$) corresponds to neutral equilibrium: a change in $|\mathfrak{M}|$ is negligible for a small change in $\delta\theta$.

Qualitatively, the HFMO excitation proceeds as follows. In the equilibrium state, $\mathbf{M}^* \parallel \mathbf{B}_{eff}$. If a magnetic moment component perpendicular to $\mathbf{B}_{eff}$ fluctuatively appears in ferromagnet F, i.e., if $[\mathbf{M}^* \times \mathbf{B}_{eff}] \neq 0$, $\mathbf{M}^*$ precesses around $\mathbf{B}_{eff}$ in the absence of damping and current. In the presence of dissipation but absence of current, this precession decays. However, in the presence of current, the current-induced magnetization flux into the ferromagnet arises if $[\mathbf{M}^* \times \mathbf{B}_{eff}] \neq 0$. This is equivalent to the appearance of an effective torque $\mathfrak{M}$ acting on $\mathbf{M}^*$. It is important that $\mathfrak{M}$ is antiparallel to dissipative torque (1) so that their sum goes to zero at a certain current magnitude and condition (3) becomes fulfilled. The viscous torque diminishes the angle θ between $\mathbf{M}^*$ and $\mathbf{B}_{eff}$, whereas $\mathfrak{M}$ increases θ. If condition (3) is met, $\mathbf{M}^*$ executes stationary precession. In particular, the role of $\mathfrak{M}$ in the HFMO excitation amounts to the compensation for viscous torque. The bell-shaped curve for the θ dependence of $|\mathfrak{M}|$ (Fig. 2b) is also essential: it has a maximum at $\theta = \theta^*$ and zero points at $\theta = 0$ and π [1, 2]. Equation (3) can be satisfied for two θ values: θ_1 and θ_2 such that $\theta_1 < \theta^* < \theta_2$. The $\theta = \theta_1$ state corresponds to unstable equilibrium, whereas the $\theta = \theta_2$ state corresponds to stable equilibrium. The character of equilibrium approach (oscillatory or aperiodic) and the possibility of the stationary state occurring are governed by the parameters of the system, in particular, by the $\mathfrak{M}(\theta)$ dependence. The state with $\theta = \theta^*$ is also noteworthy because it may be considered as a neutral equilibrium.

Within the framework of the model considered, the HFMO excitation in a disk multilayer composed of thin and thick ferromagnetic disks separated by a nonmagnetic metal [7] is described in a similar manner, with the only difference that the role of excited ferromagnet belongs to the thin ferromagnetic disk. Apart from the excited ferromagnet, it is necessary for the HFMOs that there be a "bulk" ferromagnet, in which the magnetization fluctuations are small and due to which the electrons flowing into the excited ferromagnet from the nonmagnetic metal are polarized. In the film structure, this function is fulfilled by the near-microcontact ferromagnet area of size on the order of spin-relaxation length. In the disk structure, the thick ferromagnetic disk plays the same role. Notice that neither the direction of the external magnetic field nor its magnitude is of crucial importance.

The physical nature of the threshold character of HFMO excitation is apparent. Since the magnon spectrum has a gap $\Delta_m(B_{eff})$, condition I is fulfilled only if current I exceeds a certain threshold value I^*. The I^* value is determined from the condition $\Delta_A(I^*)$ or $\Delta\varepsilon(I^*) > \Delta_m(B_{eff})$ (Fig. 1). Both mechanisms yield close I^* values for the experimental situations considered above. The current is also responsible for the appearance of a torque that sustains the fluctuative magnon generation by compensating for the dissipative losses of generation. It also provides high probability for the ferromagnet to occur in the excited state [in particular, due to $\Delta_A(I)$ and thermal heating].

There are several reasons why the HFMOs are excited only for one current direction. In the case of microcontact excitation, a high current density occurs only in the vicinity of the microcontact. For the diffusion regime of electrical conductance in the presence of ferromagnetic layer ordering and for layer thicknesses much smaller than Λ_s, the multilayer can be considered as a certain effective ferromagnet. The gap $\Delta_A(I) > \Delta_m(B_{eff})$ appears in the vicinity of microcontact for both I directions [12, 20, 21]. However, a change in the electron polarization accompanied by energy release Δ_A reduces the magnetization and provides the conditions for HFMO excitation only for one direction of I (Fig. 1a). The same reason holds for the HFMO excitation in a single-layer ferromagnet [5]. The generation of high-energy excitations [$\Delta\varepsilon(I) > \Delta_m(B_{eff})$] is insensi-

tive to the current direction (Figs. 1b, 1c). However, for the ballistic regime of electron passage through the microcontact (Fig. 1c) and flow into the ferromagnet, this generation occurs in the ferromagnet, where magnons can be created, whereas for the opposite current direction the generation proceeds in the nonmagnetic metal, where the exchange interaction is absent and the conditions for magnon excitation are cardinally different (see Supplement in [10]) and not fulfilled in [3–7]. This mechanism is also operative for the HFMO excitation in a single-layer ferromagnet [5].

The fact that the HFMOs are excited in a film magnetic multilayer with a microcontact mounted on it suggests that the real balance of all driving forces and energy conditions is positive for the HFMOs excitation. This balance is positive [3, 4, 6] if the indirect antiferromagnetic exchange interaction between the magnetic layers dominates [22], if the magnetic interaction between the magnetic layers dominates [22] when the thickness of the nonmagnetic layer is large [5, 7], and if the HFMOs are excited in a single magnetic layer [5]. However, the HFMOs were not excited upon mounting a tip on a bulk ferromagnet [23]. This is likely caused by the greater suppression effect of exchange interaction, whose magnitude is proportional to the surface area between the ferromagnet and the excited ferromagnet. For a bulk ferromagnet, this area is considerably larger (Fig. 2a). The microcontact situation is different from that considered in [24], where the current density was assumed to be uniform in the plane parallel to the N/F interface and where the authors predicted the stationarity breakdown, appearance of instability, and the possibility of HFMO excitation by a spin-polarized electron flow incident on the half-space occupied by a ferromagnet. When applying the fluctuation model to the description of a current-driven HFMO excitation, account should be taken of the results of a detailed analysis of magnetic dynamics [8] and the θ dependence of electron-assisted magnetization transfer [1].

We believe that the spin-density fluctuations; the increase in their amplitude due to the Aronov mechanism, high current density, and thermal heating; and the increased role of spin fluctuations in mesoscopic samples are precisely the factors which govern the effect of current-driven magnon generation and create conditions for the current-induced "spin-system buildup." This model fits naturally into the theory of magnetism in itinerant electron magnetics, where the spatial fluctuations of electron spin density are taken into account [25]. With the model suggested, the current-driven magnon excitation seems to be an efficient method for studying fundamentals of this theory; in particular, information can be gained about the space spin-density correlation function.

REFERENCES

1. J. Slonczewski, J. Magn. Magn. Mater. **159**, L1 (1996).
2. L. Berger, Phys. Rev. B **54**, 9353 (1996).
3. M. Tsoi, A. G. M. Jansen, J. Bass, *et al.*, Phys. Rev. Lett. **80**, 4281 (1998).
4. M. Tsoi, A. G. M. Jansen, J. Bass, *et al.*, Nature **406**, 46 (2000).
5. E. B. Myers, D. C. Ralph, J. A. Katin, *et al.*, Science **258**, 867 (1999).
6. S. M. Rezende, F. M. de Aguiar, M. A. Lucena, *et al.*, Phys. Rev. Lett. **84**, 4212 (2000).
7. J. A. Katine, F. J. Albert, R. A. Buhrman, *et al.*, Phys. Rev. Lett. **84**, 3149 (2000).
8. J. Slonczewski, J. Magn. Magn. Mater. **195**, L261 (1999).
9. J. Fabian and S. Das Sarma, cond-mat/9901170 (1999).
10. A. I. Akhiezer, V. G. Bar'yakhtar, and S. V. Peletminskii, *Spin Waves* (Nauka, Moscow, 1967; North-Holland, Amsterdam, 1968).
11. L. D. Landau and E. M. Lifshits, Phys. Z. Sowjetunion **8**, 153 (1935).
12. A. G. Aronov, Pis'ma Zh. Éksp. Teor. Fiz. **24**, 37 (1976) [JETP Lett. **24**, 32 (1976)].
13. Yu. V. Sharvin, Zh. Éksp. Teor. Fiz. **48**, 984 (1965) [Sov. Phys. JETP **21**, 655 (1965)].
14. G. Wexler, Proc. Phys. Soc. London **89**, 927 (1966).
15. N. F. Mott, Proc. R. Soc. London, Ser. A **153**, 699 (1936); **156**, 368 (1936); Adv. Phys. **13**, 325 (1964).
16. M. V. Tsoĭ, V. S. Tsoĭ, and P. Wyder, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 835 (1996) [JETP Lett. **64**, 891 (1996)].
17. V. S. Tsoi, Y. de Wilde, T. Noller, *et al.*, Europhys. Lett. **35**, 43 (1996).
18. V. S. Tsoi, J. Bass, and P. Wyder, Rev. Mod. Phys. **71**, 1641 (1999).
19. T. L. Gilbert, Phys. Rev. **100**, 1243 (1955).
20. T. Valet and A. Fert, Phys. Rev. B **48**, 7099 (1993).
21. M. Tsoi, Ph.D. Thesis (Grenoble High Magnetic Field Laboratory, 1998).
22. S. S. P. Parkin, R. Bhadra, and K. P. Roche, Phys. Rev. Lett. **66**, 2152 (1991).
23. I. K. Yanson and A. V. Khotkevich, *Atlas of Microcontact Spectra of Electron–Phonon Interaction in Metals, Handbook* (Naukova Dumka, Kiev, 1986).
24. Y. B. Bazaliy, B. A. Jones, and S.-C. Zhang, Phys. Rev. B **57**, R3213 (1998).
25. T. Moriya, *Spin Fluctuations in Itinerant Electron Magnetism* (Springer-Verlag, Berlin, 1985; Mir, Moscow, 1988).

Translated by V. Sakun