On Nambu monopole dynamics in a $SU(2)$ lattice Higgs model

M. N. Chernodub[a)]
Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia

(Submitted 1 October 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 577–580 (10 November 1997)

It is shown that an $SU(2)$ Higgs model on a lattice is equivalent to the Georgi–Glashow model in the limit of a small coupling constant between the Higgs and gauge fields. It can therefore be concluded that the transition between the confinement and symmetric phases in a 3 + 1 dimensional $SU(2)$ Higgs model at finite temperature is accompanied by condensation of Nambu monopoles. © *1997 American Institute of Physics.* [S0021-3640(97)00121-7]

PACS numbers: 11.15.Ha, 12.60.Fr, 14.80.Hv

According to one of the most popular modern hypotheses, the baryon asymmetry of the universe arose in the process of an electroweak phase transition (see, for example, the review by Rubakov and Shaposhinkoov[1]). On account of the smallness of the Weinberg angle θ_W and the insignificance of the fermion effects, this transition is largely determined by the properties of the $SU(2)$ Higgs model. The present letter examines the behavior of the magnetic fluctuations, which can play an important role in a temperature-induced phase transition, in the $SU(2)$ Higgs model.

Let us consider the $SU(2)$ lattice Higgs model with scalar field Φ_x in the fundamental representation, the action in which is described by the formula

$$S[U,\Phi]=-\frac{\beta}{2}\sum_P \mathrm{Tr}\ U_P-\frac{\kappa}{2}\sum_x\sum_\mu\ (\Phi_x^+ U_{x,\mu}\Phi_{x+\hat{\mu}}+\text{c.c.})+V(|\Phi|). \tag{1}$$

Here U_P represents the ordered product of the edge elements of the gauge field $U_{x,\mu}$ over the boundaries faces of the plaquette P, and $V(|\Phi|)$ is the potential on the field Φ, and $|\Phi|^2=\Phi^+\Phi$.

On account of the triviality of the homotopy group $\pi_2(SU(2))$, there are no topologically stable monopolar defects in this theory. However, ''embedded''[2] monopoles, the so-called ''Nambu monopoles,''[3] which are not topologically stable defects, do exist in the theory. These objects are described by the composite field

$$\chi_x^a=\Phi_x^+\sigma^a\Phi_x \tag{2}$$

(σ^a are Pauli matrices), which behaves under gauge transformations as a scalar field in the adjoint representation. A Nambu monopole is a configuration of fields U and Φ such that the field U and the composite field χ, expressed in terms of the fundamental field Φ

according to Eq. (2), possess the configuration of a 't Hooft–Polyakov monopole[4] in the Georgi–Glashow model[5] with the field χ in the adjoint representation and with the gauge field U.

Since Nambu monopoles are described solely by the gauge field U and the composite field χ, the dynamics of these monopoles is determined completely by the effective action $S_{\rm eff}$

$$e^{-S_{\rm eff}[U,\chi]}=\int \mathcal{D}\Phi\, e^{-S[U,\Phi]}\prod_a \prod_x \delta(\chi_x^a-\Phi_x^+\sigma^a\Phi_x). \tag{3}$$

To calculate the action $S_{\rm eff}$ it is convenient to study the following parametrization of the field Φ:

$$\Phi=e^{i\varphi}\Psi, \quad \Psi=\rho\begin{pmatrix}\cos\,\alpha\; e^{i\theta}\\ \sin\,\alpha\end{pmatrix},$$

where $\varphi,\theta\in[-\pi,\pi)$, $\alpha\in[0,\pi/2]$, and $\rho\in[0,+\infty)$. The fields ρ, α, and θ can be expressed in terms of the field χ^a with the aid of Eqs. (2):

$$\theta=\arctan\frac{\chi^2}{\chi^1}, \quad \alpha=\frac{1}{2}\arctan\frac{\sqrt{(\chi^1)^2+(\chi^2)^2}}{|\chi^3|}, \quad \rho=\sqrt{|\chi|},$$

whence

$$\Psi=\frac{1}{\sqrt{2\,(|\chi|-|\chi^3|)}}\begin{pmatrix}\chi^1+i\chi^2\\ |\chi|-|\chi^3|\end{pmatrix}. \tag{4}$$

Using the relation for the modulus of the field Φ, $|\Phi|^2=|\chi|=\sqrt{(\Sigma_{a=1}^3(\chi^a)^2)}$, and the measure

$$\int_{-\infty}^{+\infty}\mathcal{D}\Phi\cdots=\int_{-\pi}^{\pi}\mathcal{D}\varphi\int_{-\infty}^{+\infty}\prod_x\frac{1}{|\chi_x|}\,d\chi_x^1\,d\chi_x^2\,d\chi_x^3\cdots,$$

we obtain for the effective action (3)

$$S_{\rm eff}[U,\chi]=-\frac{\beta}{2}\sum_P \mathrm{Tr}\; U_P+S_h[U,\chi]+\widetilde{V}(|\chi|), \tag{5}$$

where the new potential on the field χ is determined by the expression

$$\widetilde{V}(|\chi|)=V(\sqrt{|\chi|})+\sum_x \ln|\chi_x|, \tag{6}$$

and the interaction of the fields U and χ is

$$e^{-S_h[U,\chi]}=\int_{-\pi}^{\pi}\mathcal{D}\varphi\;\exp\left\{\kappa\sum_x\sum_\mu R_{x,\mu}\cos(\varphi_{x+\hat{\mu}}-\varphi_x+A_{x,\mu})\right\}. \tag{7}$$

In this formula we introduced the notation

$$\Psi_x^+U_{x,\mu}\Psi_{x+\hat{\mu}}=R_{x,\mu}e^{iA_{x,\mu}}. \tag{8}$$

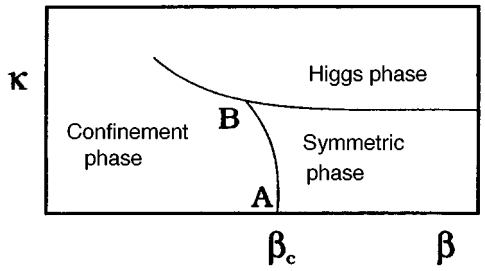


FIG. 1. Phase diagram of the $SU(2)$ Higgs model for small values of κ.

The derivation of the effective action (5) is correct in any dimension of space–time.

For simplicity, we shall examine the case of an infinitely deep potential $V(|\Phi|)$ with a minimum at $|\Phi|^2=|\chi|=1$. In this case, the lengths of the Higgs field Φ and of the composite field χ are frozen. The integral (7) is most easily calculated in the limit $\kappa \ll 1$. In leading order we obtain (to within a constant)

$$S_h=-\frac{\kappa^2}{2}\sum_x \sum_\mu R^2_{x,\mu}+O(\kappa^4)=-\frac{\kappa^2}{8}\sum_x \sum_\mu \mathrm{Tr}(U_{x,\mu}\chi_x U^+_{x,\mu}\chi_{x+\hat{\mu}})+O(\kappa^4),$$

where we employed Eqs. (4) and (16) and introduced the notation $\chi=\chi^a\sigma^a$. Thus in the limit $\kappa \ll 1$ the effective action (5) with the length $|\Phi|^2=1$ of the Higgs field frozen is identical in leading order to the Georgi–Glashow action

$$S_{\mathrm{eff}}[U,\chi]=-\frac{\beta}{2}\sum_P \mathrm{Tr}\ U_P-\frac{\gamma}{2}\sum_x \sum_\mu \mathrm{Tr}(U_{x,\mu}\chi_x U^+_{x,\mu}\chi_{x+\hat{\mu}})+O(\kappa^4), \tag{9}$$

where

$$\gamma=\kappa^2/4\,. \tag{10}$$

It is interesting to compare the phase diagrams of the 3 + 1-dimensional $SU(2)$ Higgs model (1) and the Georgi–Glashow model (9)–(10) at nonzero temperature for small values of the parameter κ (γ, respectively). Figure 1 displays schematically the phase diagram obtained in Ref. 7 for the $SU(2)$ Higgs model. Figure 2 shows the phase

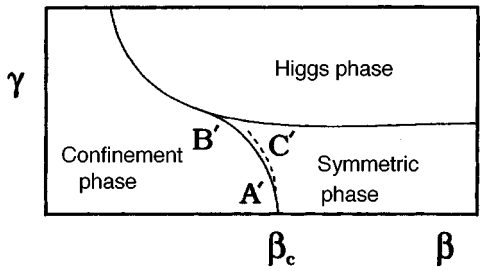


FIG. 2. Phase diagram of the Georgi–Glashow model for small γ.

diagram obtained in Ref. 8 for the Georgi–Glashow model. For small values of the constant β both theories are in the confinement phase (color confinement). As β increases, a phase transition from the confinement phase to the symmetric phase occurs in both theories, and the line of phase transitions $A'-B'$ in the Georgi–Glashow model should correspond to the line of phase transitions $A-B$ in the $SU(2)$ Higgs model according to Eq. (10): $\gamma_c(\beta)=\kappa_c^2(\beta)/4+O(\kappa_c^4)$. Figure 2 shows schematically the phase transition predicted with the aid of Eq. (10) (dashed line $A'-C'$). Unfortunately, it is impossible to determine the correctness of this prediction quantitatively on the basis of the results of Refs. 7 and 8, since in those papers the phase diagrams were studied at different temperatures.

It is known[6] that in the Georgi–Glashow model confinement is due to the dynamics of the 't Hooft–Polyakov monopoles: in the confinement phase the monopoles are condensed, while in the deconfinement phase there exists a dilute gas of monopole–antimonopole pairs. Therefore, at least for low values of the coupling constant κ, the phase tranistion from the symmetric phase to the confinement phase in the $SU(2)$ Higgs model (1) is accompanied by condensation of Nambu monopoles, since the Nambu monopoles in the $SU(2)$ Higgs model (1) correspond to the 't Hooft–Polyakov monopoles in the Georgi–Glashow model (9).

It is natural to suppose that condensation of Nambu monopoles also occurs for larger values of the parameter κ in the course of the phase transitions from the confinement phase to the symmetric phase and from the confinement phase to the Higgs phase. The latter conjecture finds support in the fact that in the Higgs phase there exists an embedded string[3,9] with nonzero string tension. Stretched between Nambu monopoles, such a string destroys the condensate. The results of investigations of this question will be published later in a separate paper.

I thank E.-M. Ilgenfritz and M. I. Polikarpov for helpful remarks. This work was supported in part by the Russian Fund for Fundamental Research under Grant No. 96-02-17230a and under the Grants INTAS-94-0840 and INTAS-RFBR-95-0681. I also thank the Japan Society for the Promotion of Science (JSPS) for financial assistance, provided as part of the program for the support of scientists of the former Soviet Union.

[a)]e-mail: chernodub@vxitep.itep.ru

[1] V. A. Rubakov and M. E. Shaposhnikov, Usp. Fiz. Nauk **166**, 493 (1996).
[2] T. Vachaspati and M. Bariolla, Phys. Rev. Lett. **69**, 1867 (1992); M. Bariolla, T. Vachaspati, and M. Bucher, Phys. Rev. D **50**, 2819 (1994).
[3] Y. Nambu, Nucl. Phys. B **130**, 505 (1977).
[4] G. 't Hooft, Nucl. Phys. B **79**, 276 (1974); A. M. Polyakov, JETP Lett. **20**, 194 (1974).
[5] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **28**, 1494 (1972).
[6] A. M. Polyakov, Nucl. Phys. B **120**, 429 (1977).
[7] H. G. Evertz, J. Jersak, and K. Kanaya, Nucl. Phys. B **285**, 229 (1987).
[8] F. Karsch, E. Seiler, and I. O. Stamatescu, Phys. Lett. B **131**, 138 (1983).
[9] N. S. Manton, Phys. Rev. D **28**, 2019 (1983).

Translated by M. E. Alferieff

Giant enhancement of sum-frequency generation upon excitation of a surface plasmon–polariton

E. V. Alieva, Yu. E. Petrov, and V. A. Yakovlev[a)]
Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

E. R. Eliel, E. W. M. van der Ham, and Q. H. F. Vrehen
Huygens Laboratory, Leyden University, 2300 RA Leyden, The Netherlands

A. F. G. van der Meer
FOM-Institute for Plasma Physics Rijnhuizen, P. O. Box 1207, 3430 BE Nieuwegein, The Netherlands

V. A. Sychugov
Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia

(Submitted 18 September 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 581–584 (10 November 1997)

The generation of the sum frequency of visible (0.5235 μm) and IR (10 μm) radiation on smooth and corrugated silver surfaces is investigated. The sum-frequency signal obtained with a visible-range surface plasmon–polariton excited on a corrugated silver–air interface is found to be more than four orders of magnitude stronger than the signal from a smooth silver surface. © *1997 American Institute of Physics.* [S0021-3640(97)00221-1]

PACS numbers: 73.20.Mf, 71.36.+c

Nonlinear sum-frequency generation (SFG) spectroscopy has been under active development in recent years as a tool for investigating surfaces and interfaces.[1] SFG methods are especially promising for the investigation of the surfaces of media containing centers of inversion, since in this case nonlinear optical processes occur only on the surface. The investigation of thin films, right down to monolayer thicknesses, on the surface of these materials is also promising. However, the SFG signal from most surfaces is very weak.

The SFG signal can be intensified by resonantly increasing the strength of the electric field at the surface under study. One of the simplest and most effective methods of enhancing the field on the surface of a metal is excitation of a surface polariton.[2] A surface plasmon–polariton (SPP) can exist on metal surfaces in a wide spectral range — from the far-IR up to UV.[3] The electromagnetic field of a SPP is strongest at an interface and decreases exponentially with distance from the interface.

We employed a diffraction grating (corrugated silver surface) to excite a SPP.[4] In this case, the SPP is excited by a *p*-polarized beam of light incident at a prescribed angle onto a surface which is corrugated with a period of the order of the wavelength.

A diagram of the experiment is displayed in Fig. 1. The setup employed in the

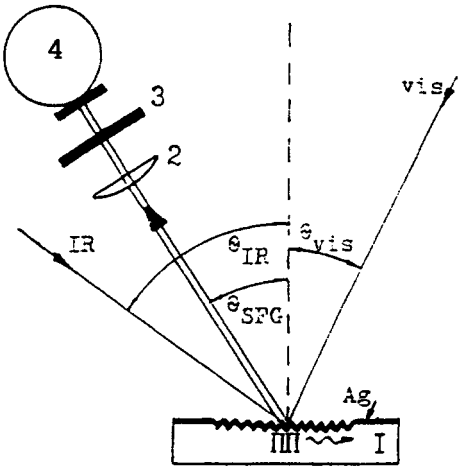


FIG. 1. Diagram of the experimental setup: *1* — Sample, *2* — focusing lens, *3* — filter, *4* — CCD camera.

present work is described in Ref. 5. The FELIX free-electron laser was used as the IR radiation source.[6] The visible-range radiation source operates at the second harmonic of a Nd:YLF laser (the wavelength of the second harmonic equals 523.5 nm). The two lasers are synchronized with each other so as to be able to obtain the maximum temporal overlap of the micropulses at the required point through a continuous adjustment of the delay of the visible radiation pulse. The energy of the FELIX micropulse and of the visible-range laser pulse equalled, on average, 10 and 8 μJ, respectively. All measurements were performed at a fixed FELIX radiation wavelength — 10 μm. The sum-frequency signal was detected with a liquid-nitrogen-cooled Princeton Instruments CCD (charge-coupled device) camera. The SFG signal was emitted from the surface in the form of an approximately parallel beam, which was focused by a lens onto the CCD camera. The focal length of the lens was equal to 20 cm. A replaceable selective narrow-band filter (transmission coefficient — 70% at the frequency of the SFG signal and 10^{-5} at the frequency of the Nd:YLF laser) was used to suppress the strong scattered radiation of the Nd:YLF laser. The signal with the IR radiation blocked off was recorded to make sure of the absence of a Nd:YLF-laser signal.

A diffraction grating was deposited on part of the surface of glass plates by a photolithographic method followed by ion etching. The period of the gratings (determined according to the diffraction of a visible-range laser beam) was equal to 305 nm for grating I and 318 nm for grating II; the depth of the gratings was 15–20 nm. A silver film approximately 200 nm thick was deposited on the grating from above by thermal evaporation in vacuum with a residual-gas pressure of $\approx 10^{-5}$ torr.

In the excitation of a SPP at an interface the tangential component of the momentum must be conserved. In the case of a diffraction grating deposited on a surface, and with both the visible-range and IR beams lying in a plane perpendicular to the grating lines, this condition for an air–metal interface can be written in the form

TABLE I. Experimental parameters and intensities (in the pulses) of the SFG signal measured with a CCD camera for different configurations of the samples.

Sample	Lattice period, nm	θ_{vis}	n'	I_{sfg}		
				config. 1	config. 2	smooth metal
I	305	40.3°	1.069	1×10^5	82	57
II	318	35.5°	1.065	1.5×10^6	–	98

$$\frac{1}{\lambda_{\text{vis}}}\sin\theta_{\text{vis}}-\frac{1}{\lambda_{ir}}\sin\theta_{ir}=\frac{1}{\lambda_{sfg}}\sin\theta_{sfg}, \tag{1}$$

where λ_{vis}, λ_{ir}, and λ_{sfg} are, respectively, the wavelengths of the visible-range laser radiation, the infrared FELIX radiation, and the generated sum-frequency radiation, and θ_{vis}, θ_{ir}, and θ_{sfg} are the corresponding angles of incidence (emission).

The same equation also holds in the case of SFG on a smooth surface of the sample.[6] Hence it follows that we can compare the SFG intensities from smooth and corrugated surfaces by simple parallel displacement of the sample, keeping the entire geometry of the experiment unchanged.

In the experiment the angles of incidence of the visible-range and IR beams on the sample were varied simultaneously by rotating the sample. The angle between them was fixed and equal to 90.3°. The angle of maximum excitation of the SPP was determined experimentally by two independent methods: according to the minimum of the specular reflection and visually according to the maximum of the light scattered by the sample. The two methods gave identical values to within the experimental error. The angles of incidence θ_{vis}, corresponding to maximum excitation of the SPP are given in Table I. The values of the real part n' of the effective refractive index of the SPP[3] calculated using these values (and also presented in Table I) are in good agreement with the value 1.056 calculated from the optical constants of silver, which are given in Ref. 7 (see also Ref. 8).

The SFG yield was measured in three different configurations:

1. With the wave vector of the visible-range radiation parallel to the dispersion plane of the grating.

2. With the wave vector of the visible-range radiation perpendicular to the dispersion plane of the grating.

3. With the sum-frequency generation occurring on the smooth surface of the metal outside the diffraction grating.

The transition between these configurations is made by rotating the sample by 90° or by parallel displacement of the sample. The results of these three measurements are also presented in Table I. As one can see from the table, in the case of excitation of a SPP (configuration 1) the enhancement of the SFG is by a factor of $\approx10^3$ for sample I and $\approx1.5\times10^4$ for sample II. The absence of enhancement in configuration 2 (a SPP is not

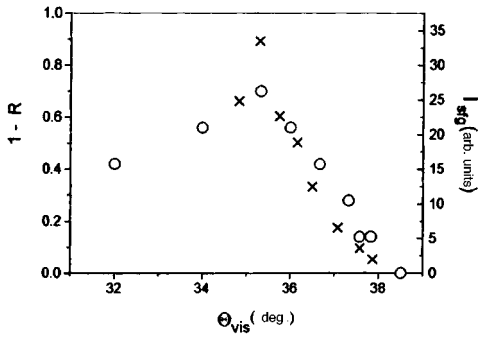


FIG. 2. $1-R$ (○) and SFG intensity I_{sfg} (×) versus the angle of incidence of the visible-range radiation for sample II in the region of SPP excitation.

excited) indicates that this effect is not associated with an increase in the roughness of the surface, which might have taken place in the course of preparation of the diffraction grating.

It is well known that the angular dependence of the intensity of the electromagnetic radiation specularly reflected from a diffraction grating has a minimum near the angle of excitation of a SPP.[3,4] If the depth of the grating is small compared with the wavelength of the incident radiation, as is the case for all the gratings employed in the present work, the intensity of the SPP and therefore also $I_{sfg}(\theta_{\rm vis})$ should be proportional to $1-R(\theta_{\rm vis})$, where $R(\theta_{\rm vis})$ is the reflectivity of the grating for visible-range radiation.[3] The curves $I_{sfg}(\theta_{\rm vis})$ and $1-R(\theta_{\rm vis})$ for sample II are presented in Fig. 2. As one can see from the figure, the indicated relation holds to within the experimental error. Although our results indicate unequivocally that the excitation of a SPP is responsible for the enhancement of SFG, the value obtained for the enhancement factor is approximately two orders of magnitude greater than the expected enhancement of the electric field of the wave of exciting radiation on the corrugated surface.[9] Further investigations are needed to determine the reasons for such a large discrepancy.

We thank the The Netherlands Fund for Research on the Fundamental Properties of Materials (FOM) for support and for providing the time required for working on FELIX. This work is part of the research program of the The Netherlands Technological Foundation (STW) and is supported by Russian Fund for Fundamental Research Grants (Nos. 95-02-04194, 95-02-04195, and 97-02-16792). We thank L. A. Kuzik and I. F. Salakhutdinov for assistance in setting up the experiment.

[a)]e-mail: yakovlev@isan.troitsk.ru

[1] Y. R. Shen, Nature **337**, 519 (1989).
[2] E. V. Alieva, G. N. Zhizhin, V. A. Yakovlev *et al.*, JETP Lett. **62**, 817 (1995).
[3] V. M. Agranovich and D. L. Mills (eds.), *Surface Polaritons*, North-Holland, Amsterdam, 1982 [Russian translation, Nauka, Moscow, 1985].
[4] D. Beaglehole, Phys. Rev. Lett. **22**, 706 (1969).
[5] E. W. M. van der Ham, Q. H. F. Vrehen, and E. R. Eliel, Opt. Lett. **21**, 1448 (1996).

[6]R. J. Bakker, D. A. Jaroszynski, A. F. G. van der Meer *et al.*, IEEE J. Quantum Electron. **QE-30**, 1635 (1994).
[7]P. B. Jonson and R. V. Christy, Phys. Rev. B **6**, 4370 (1972).
[8]M. A. Ordal, L. L. Long, R. J. Bell *et al.*, Appl. Opt. **22**, 1099 (1983).
[9]D. L. Mills and M. Weber, Phys. Rev. B **26**, 1075 (1982).

Translated by M. E. Alferieff

Photo- and conversion electron yield from a target irradiated with a ^{57}Co source at glancing incidence of the primary beam

A. S. Serebryakov and V. V. Smirnov
V. G. Khlopin Radium Institute Scientific–Industrial Association, 194021 St. Petersburg, Russia

(Submitted 18 September 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 585–587 (10 November 1997)

A Monte Carlo calculation of photo- and conversion electron yields and spectra from a Fe target shows that the yield and hardening of the spectrum increase with increasing grazing angles of incidence of the primary beam, and that this effect is sharper for photoelectrons, indicating that their ionization efficiency is higher. © *1997 American Institute of Physics.* [S0021-3640(97)00321-6]

PACS numbers: 79.60.−i, 02.70.Lq

The photoelectron yield from targets irradiated with an x-ray beam at different angles of incidence was studied experimentally and theoretically in Refs. 1 and 2. A Monte Carlo calculation showed that the anisotropy produced in the initial angular distribution of photoelectrons by a variation in the degree of polarization of the primary photon beam has virtually no effect on the integrated electron yield from the target. However, these calculations did not treat the case of glancing incidence of the primary beam on the target. Moreover, the anisotropy due to the predominant emergence of the photoelectrons in the direction of the electric field vector of the incident electromagnetic wave is small and is described by a $\cos^2\phi$ law, where ϕ is the azimuthal angle of emergence.

At the same time, for glancing incidence of a beam on a target, the substantial anisotropy of the initial distribution of the photoelectrons over the polar angle θ of emergence should result in a large change in the integrated yield. This is due to two circumstances: First, the initial angular distribution of emergence (Fisher or Sauter, depending on the energy of the photoelectron) has a sharp maximum at angles θ close to $\pi/2$, and this means that some photoelectrons move from inside the target in a direction almost normal to its surface; second, glancing incidence of the primary beam makes for photoelectron production in an ultrathin layer of the target, and therefore the trajectories of these photoelectrons before emerging from the target can retain information about their initial angular distribution. Thus it is obvious that, as compared with an isotropic initial angular distribution, some hardening of the spectrum of the emerging electrons should be observed in addition to an increase in the integrated photoelectron yield.

A model simulating the production and transport of electrons was constructed for two types of particles: photoelectrons and conversion electrons (CEs) produced in an iron target by a well-collimated beam of radiation from ^{57}Co. It was assumed that the initial

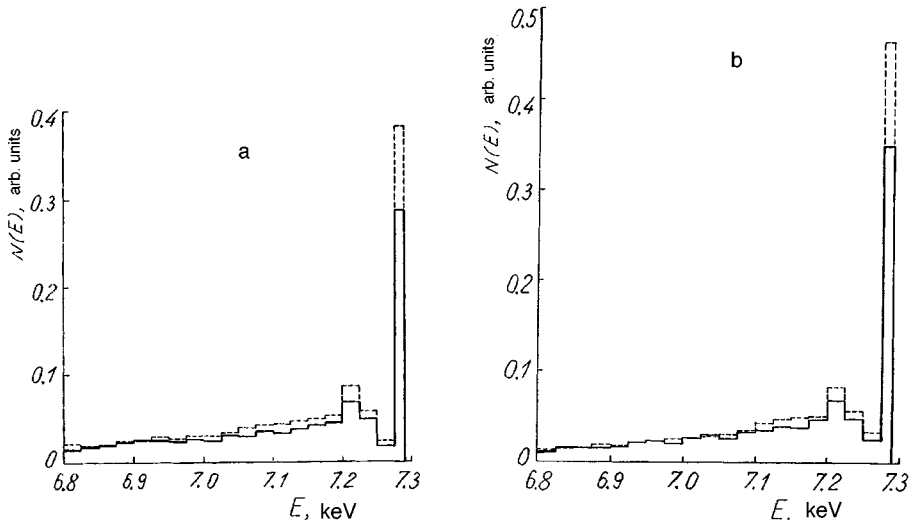


FIG. 1. Conversion-electron spectra (solid lines) and photoelectron spectra (dashed lines) from an iron target at grazing angles of incidence α of the primary beam from a ^{57}Co source: a — $\alpha = 1.754$ mrad, b — $\alpha = 0.853$ mrad.

angular distribution of the CEs is strictly isotropic. The model employed the differential elastic electron scattering cross sections tabulated in Ref. 3, the total inelastic electron scattering cross sections from Ref. 4, and the method described in Ref. 5 for modeling energy losses in inelastic collisions. Conversion-electron spectra for electrons emerging from different depths in the target, as presented in Ref. 6, were obtained in order to adjust the simulation program.

Assuming that the production probabilities of the two types of electrons are identical, the spectra of the particles emerging from the target were calculated for two grazing angles of incidence α of the primary radiation beam on the target — 1.754 and 0.853 mrad relative to the surface of the target (see Fig. 1). The total mass coefficient of absorption for the primary beam in the target material was equal to 0.0636 cm^2/mg, which corresponds to the photoabsorption cross section for the K shell of iron atoms for an energy of 14.4 keV, and the number of ''successful'' trajectories, on which the electrons emerged from the target was equal to 2×10^4 in each case.

Peaks with a maximum initial energy of 7.29 keV can be seen in the spectra. These peaks are due to electrons which have undergone only elastic collisions in the target. As the angle α increases, the integrated yields for both groups of electrons increase and the photoelectron and CE spectra harden. But this tendency is sharper in the photoelectron spectra, i.e., the number of emerging photoelectrons in the energy range 6.8–7.29 keV at fixed angles of incidence of the primary beam on the target is larger by 10.5 and 8.8%, respectively, than the number of emerging CEs, and their spectrum is harder. Each spectrum in the figure is normalized to the total number of emerging electrons of a given type. Therefore the total ionization effect produced by electrons in the detector (for

example, in a gas proportional counter) for the same initial number of particles of both types is appreciably larger for photoelectrons than for CEs.

In our opinion, these results explain the distortion of the information-carrying signal in experiments on the Mössbauer spectroscopy of ultrathin layers of matter at glancing incidence of the radiation beam on target.[7] In this case it is assumed that the signal in the detector is entirely due to the emerging CEs. However, photoelectrons are produced in the target together with CEs, and the contribution of the photoelectrons to the information-carrying signal is found to be substantial. Specifically, the observed rise in the amplitudes on the wings of the resonance absorption curve, where the cross section for resonance absorption of photons drops to and below the level of the photoabsorption cross section, is attributed to the increasing contribution of photoelectrons in the energy interval under study.

[1] M. V. Koval'chuk, D. Dil'ekvist, and V. G. Kon, Fiz. Tverd. Tela (Leningrad) **28,** 3409 (1986) [Sov. Phys. Solid State **28**, 1918 (1986)].
[2] M. V. Kovalchuk and D. Liljequist, USIP Report 86-03, Stockholm, 1986.
[3] M. E. Riley, C. J. McCallum, and F. Biggs, At. Data Nucl. Data Tables **15**, 443 (1975).
[4] A. F. Akkerman, *Simulation of Charged Particle Trajectories in Matter* [in Russian], Énergoatomizdat, Moscow, 1991.
[5] F. Salvat and J. Parellada, J. Phys. D: Appl. Phys. **17**, 185 (1984).
[6] D. Liljequist, T. Ekdahl, and U. Baverstam, Nucl. Instrum. Methods **155**, 529 (1978).
[7] M. A. Andreev, S. M. Irkaev, and V. G. Semenov, Zh. Éksp. Teor. Fiz. **105**, 1767 (1994) [JETP **78**, 956 (1994)].

Translated by M. E. Alferieff

Intraexcitonic transitions in two-dimensional systems in a high magnetic field

A. B. Dzyubenko
Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia

(Submitted 30 September 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 588–593 (10 November 1997)

The internal transitions of two-dimensional (2D) excitons in a high magnetic field B exhibit features due to the coupling of the internal and center-of-mass motions. A study is made of these features, and it is shown that for magnetoexcitons with a center-of-mass momentum $\mathbf{K} \neq 0$ the energies of the strong transitions decrease with increasing $\mathbf{K}$, and the absorption spectra show weakly resolved transitions, whose total intensity depends strongly on the exciton statistics (distribution function). © *1997 American Institute of Physics.*
[S0021-3640(97)00421-0]

PACS numbers: 71.35.Ji, 75.70.Cn

1. Intraband transitions of quasi–two-dimensional excitons in quantum wells (QWs) and superlattices in a magnetic field have attracted a great deal of interest in recent years (see Refs. 1–3 and the literature cited therein). Progress in this field requires a sensitive method of investigation — optically detected cyclotron resonance. Intraband IR magneto-spectroscopy could be effective for studying the kinetics of interlevel excitonic transitions, for investigating collective effects in a system of excitons with finite density, and for resolving the fine structure of the ground and excited states of quasi-2D excitons, for example, in coupled double QWs.[4]

In the case of intraband IR spectroscopy, all populated excitonic states give a response, including states with finite center-of-mass momentum $\mathbf{K}$. This is in contrast to interband transitions for which only excitons with $\mathbf{K}=0$ are optically active. Physically, the center-of-mass and relative motions of a neutral e–h pair are coupled in a magnetic field B. The present letter examines theoretically some characteristics of excitonic IR absorption, which are associated with this circumstance, in 2D systems in a high magnetic field. Similar effects should exist in atomic physics (taking account of the change in the characteristic magnetic field and momentum scales[5]).

2. For simplicity, we shall study the purely 2D situation. Motion of a 2D neutral e–h pair in a transverse magnetic field $\mathbf{B}=(0,0,B)$ is described by the Hamiltonian

$$H=\frac{1}{2m_e}\left(-i\hbar\nabla_e+\frac{e}{c}\mathbf{A}_e\right)^2+\frac{1}{2m_h}\left(-i\hbar\nabla_h-\frac{e}{c}\mathbf{A}_h\right)^2-\frac{e^2}{\epsilon|\mathbf{r}_e-\mathbf{r}_h|}\equiv H_0+U_{eh}, \quad (1)$$

where $\mathbf{r}=(x,y)$. The motion is characterized[5] by a conserved magnetic momentum of the center of mass $\hat{\mathbf{K}}=-i\hbar\nabla_{\mathbf{R}}-(e/c)\,\mathbf{A}(\mathbf{r})$. Here $\mathbf{R}=(m_e\mathbf{r}_e+m_h\mathbf{r}_h)/M$ are the coordinates of the center of mass and $\mathbf{r}=\mathbf{r}_e-\mathbf{r}_h$ are the relative e–h coordinates, $M=m_e+m_h$, and

$\mathbf{A}=\frac{1}{2}\mathbf{B}\times\mathbf{r}$. The wave function of an exciton with momentum $\mathbf{K}$ can be represented in the form[5] $\Psi_{\mathbf{K}}=\exp(i/\hbar)[\mathbf{K}+(e/c)\,\mathbf{A}(\mathbf{r})]\cdot\mathbf{R})\Phi_{\mathbf{K}}(\mathbf{r})$. This can also be regarded as a unitary transformation of the Hamiltonian $H\to\tilde{H}(\mathbf{K})=\hat{U}^{\dagger}H\hat{U}$, where $\hat{U}(\mathbf{K})=\exp((i/\hbar)\times[\mathbf{K}+(e/c)\,\mathbf{A}(\mathbf{r})]\cdot\mathbf{R})$. The transformed Hamiltonian has the form[5,6] $\tilde{H}(\mathbf{K})=\tilde{H}_0(\mathbf{K})+U_{eh}$, and

$$\tilde{H}_0(\mathbf{K})=-\frac{\hbar^2}{2\mu}\nabla_{\mathbf{r}}^2+\frac{1}{2}\hbar(\omega_{ch}-\omega_{ce})\hat{l}_z+\frac{e^2B^2}{8\mu c^2}r^2+\frac{e}{Mc}\mathbf{B}\cdot[\mathbf{r}\times\mathbf{K}]+\frac{K^2}{2M}, \tag{2}$$

where $\mu^{-1}=m_e^{-1}+m_h^{-1}$, $\omega_{ce(h)}=eB/m_{e(h)}c$, and $\hat{l}_z=-i[\mathbf{r}\times\nabla_{\mathbf{r}}]_z$ is the projection of the angular momentum of the relative motion. Similarly to the case of electrons in a magnetic field B, the Hamiltonian (2) can be diagonalized in a representation of Bose ladder operators (see Ref. 7). For this, we first perform another unitary transformation[5] $\tilde{H}(\mathbf{K})\to\overline{H}=\hat{W}^{\dagger}(\mathbf{K})\tilde{H}(\mathbf{K})\hat{W}(\mathbf{K})$, where $\hat{W}(\mathbf{K})=\exp((i/2\hbar)\,\gamma\mathbf{K}\cdot\mathbf{r})$ and $\gamma=(m_h-m_e)/M$, and then a translation of the coordinates $\mathbf{r}\to\overline{\mathbf{r}}=\mathbf{r}-\mathbf{r}_0=(\overline{x},\overline{y})$ with $\mathbf{r}_0=\mathbf{e}_z\times\mathbf{K}l_B^2/\hbar$. After this we obtain the Hamiltonian $\overline{H_0}$ (obviously, $\overline{H}=\overline{H_0}+U_{eh}(\overline{\mathbf{r}})$) which in the coordinate representation assumes the form of the Hamiltonian $\tilde{H}_0(\mathbf{K}=0)$ from Eq. (2). To diagonalize $\overline{H_0}$ we introduce the ladder operators

$$\overline{a^{\dagger}}=\frac{1}{\sqrt{2}}\left(\frac{z}{2l_B}-2l_B\frac{\partial}{\partial z^*}\right),\quad \overline{b^{\dagger}}=\frac{1}{\sqrt{2}}\left(\frac{z^*}{2l_B}-2l_B\frac{\partial}{\partial z}\right), \tag{3}$$

such that $[\overline{a},\overline{a^{\dagger}}]=[\overline{b},\overline{b^{\dagger}}]=1$ and $[\overline{a},\overline{b}]=[\overline{a},\overline{b^{\dagger}}]=0$; here $z=\overline{x}+i\overline{y}$ and $l_B=(\hbar c/eB)^{1/2}$. In this representation, we have $\overline{H_0}=\hbar\omega_{ce}(\overline{a^{\dagger}}\,\overline{a}+\frac{1}{2})+\hbar\omega_{ch}(\overline{b^{\dagger}}\,\overline{b}+\frac{1}{2})$, so that the orthonormalized eigenstates have the form of factorized wave functions $|nm\rangle=(\overline{a^{\dagger}})^n(\overline{b^{\dagger}})^m|00\rangle/\sqrt{n!m!}$ with eigenvalues $\hbar\omega_{ce}(n+\frac{1}{2})+\hbar\omega_{ch}(m+\frac{1}{2})$. In the coordinate representation the wave functions $\langle\mathbf{r}|nm\rangle\equiv\phi_{nm}(\mathbf{r})$ are identical to the wave functions of an electron in a field B (for example, $\langle\mathbf{r}|\overline{00}\rangle=\exp(-\rho^2/4l_B^2)/(2\pi l_B^2)^{1/2}$). In the case of a magnetoexciton the operators $\overline{a^{\dagger}}$, $\overline{a}$ ($\overline{b^{\dagger}}$, $\overline{b}$) describe electronic (hole) Landau levels. Since $\hat{S}(\mathbf{K})\equiv\hat{W}(\mathbf{K})\hat{U}(\mathbf{K})=\exp((i/\hbar)\mathbf{R}_0\cdot[\mathbf{K}+(e/c)\,\mathbf{A}(\mathbf{r})])$, where $\mathbf{R}_0=\frac{1}{2}(\mathbf{r}_e+\mathbf{r}_h)$, the wave functions $|nm\mathbf{K}\rangle=\hat{S}(\mathbf{K})|nm\rangle$ describing the free motion of an e–h pair in a field B can be represented in the form

$$\Psi_{nm\mathbf{K}}(\mathbf{r}_e,\mathbf{r}_h)=\langle\mathbf{r}_e\mathbf{r}_h|nm\mathbf{K}\rangle=\exp\left(\frac{i}{\hbar}\mathbf{R}_0\cdot\left[\mathbf{K}+\frac{e}{c}\mathbf{A}(\mathbf{r})\right]\right)\phi_{nm}(\mathbf{r}-\mathbf{r}_0). \tag{4}$$

The wave functions $\Psi_{nm\mathbf{K}}(\mathbf{r}_e,\mathbf{r}_h)$ correspond in the limit of a high magnetic field (cf. Ref. 6) to 2D magnetoexcitons with the dispersion relation $E_{nm}(\mathbf{K})=\langle nm|U_{eh}(\mathbf{r}-\mathbf{r}_0)|nm\rangle$.

3. Let us examine the interaction of excitons with IR radiation. In the Faraday geometry (the radiation propagates parallel to **B**) the Hamiltonian describing absorption accompanying an interaction with the ac electric field (with amplitude $\mathcal{F}_0$ and frequency ω) of circularly polarized IR radiation has the form

$$\delta\hat{V}^{\pm}=\frac{e\mathcal{F}_0}{\omega}\left(\frac{\pi_e^{\pm}}{m_e}-\frac{\pi_h^{\pm}}{m_h}\right)\exp(-i\omega t). \tag{5}$$

Here the $\pm$ signs denote left (right) circular polarization $\sigma^{\pm}$, and

$$\pi_j^{\pm}=\pi_{jx}\pm i\pi_{jy},(j=e,h), \quad \boldsymbol{\pi}_e=-i\hbar\boldsymbol{\nabla}_e+\frac{e}{c}\mathbf{A}_e, \quad \boldsymbol{\pi}_h=-i\hbar\boldsymbol{\nabla}_h-\frac{e}{c}\mathbf{A}_h.$$

It can be shown that $[\delta\hat{V}^{\pm},\hat{\mathbf{K}}]=0$, i.e., magnetic momentum is conserved in IR transitions (in the dipole approximation this also follows from the law of conservation of the total momentum). For $\mathbf{K}=0$, magnetoexcitons can be characterized by the conserved projection of the angular momentum l_z of the relative e–h motion; here $l_z=n-m$ ($\hat{l}_z=\overline{a^{\dagger}}\,\overline{a}-\overline{b^{\dagger}}\,\overline{b}$). For this reason, for excitons with $\mathbf{K}=0$ in a field B the selection rules have the standard form

$$\langle\Psi'_{\mathbf{K}=0,l'_z}|\delta\hat{V}^{\pm}|\Psi_{\mathbf{K}=0,l_z}\rangle\sim\delta_{l'_z,l_z\pm1}. \tag{6}$$

For $\mathbf{K}\neq0$, on account of the presence of the term $(e/Mc)\mathbf{B}\cdot[\mathbf{r}\times\mathbf{K}]$ (which corresponds to a uniform electric field in the moving coordinate system in B), the Hamiltonian (2) does not possess axial symmetry. As a result, the selection rules for IR transitions reduce to only conservation of momentum: generally speaking, $\langle\Psi'_{\mathbf{K}}|\delta\hat{V}^{\pm}|\Psi_{\mathbf{K}}\rangle\neq0$ for all pairs of excitonic terms. The analysis simplifies in the high-field limit. The matrix elements of the operator describing the interaction with the IR radiation field between states of the 2D magnetoexcitons (4) have the form

$$\langle n'm'\mathbf{K}|\delta\hat{V}^{\pm}|nm\mathbf{K}\rangle=\langle n'm'|\hat{S}(\mathbf{K})^{\dagger}\delta\hat{V}^{\pm}\hat{S}(\mathbf{K})|nm\rangle. \tag{7}$$

The relation

$$\hat{S}(\mathbf{K})^{\dagger}\delta\hat{V}^{+}\hat{S}(\mathbf{K})=\frac{i\sqrt{2}e\hbar\mathcal{F}_0}{\omega l_B}\left(\frac{\overline{a^{\dagger}}}{m_e}-\frac{\overline{b}}{m_h}\right)e^{-i\omega t} \tag{8}$$

shows that the matrix element (7) does not depend on the momentum $\mathbf{K}$, and in this limit transitions are possible only with a change in the Landau level numbers $\Delta n(\Delta m)=1$ for $\sigma^{\pm}$ polarization. The mixing of the Landau levels is taken into account below.

4. Let us consider IR transitions between excitons with $\mathbf{K}=0$. In high magnetic fields, the $1s$ excitonic states are formed mainly by the state $|00\mathbf{K}=0\rangle$, which corresponds to the zeroth e and h Landau levels. On account of the e–h Coulomb interaction, there is also a weak $\sim l_B/a_{Be(h)}\ll1$ [$a_{Be(h)}=\epsilon\hbar^2/m_{e(h)}e^2$] admixing of higher Landau levels $|nn\mathbf{K}=0\rangle$. In a similar manner, the $2p^+(2p^-)$ excitonic states are formed in the ground state $|10\mathbf{K}=0\rangle$ ($|01\mathbf{K}=0\rangle$) with a weak admixing of the states $|n+1\,n\mathbf{K}=0\rangle$ ($|nn+1\,\mathbf{K}=0\rangle$). For this reason, the excitonic transition $1s\rightarrow2p^+$ ($1s\rightarrow2p^-$) in a high field B can be regarded[4] as an electron (hole) cyclotron resonance $\phi_{00}\rightarrow\phi_{10}$ ($\phi_{00}\rightarrow\phi_{01}$), which is modified by excitonic effects. In the purely 2D case and in the limit of a high magnetic field, the binding energies of $1s$ and $2p^{\pm}$ magnetoexcitons are equal to[6] $E_{00}=E_0$ and $E_{10}=E_{01}=\frac{1}{2}E_0$, respectively; here $E_0=\sqrt{\pi/2}e^2/\epsilon l_B\sim\sqrt{B}$. For this reason, the $1s\rightarrow2p^{\pm}$ transition energies in this limit are

$$E_{1s\to 2p^+}=\hbar\omega_{ce}+\frac{1}{2}E_0,\quad E_{1s\to 2p^-}=\hbar\omega_{ch}+\frac{1}{2}E_0. \tag{9}$$

The transitions $1s\to np^{\pm}$ to higher-lying excited states are weak $\sim[l_B/a_{Be(\mathrm{h})}]^2$, and their energies

$$E_{1s\to np^{\pm}}=\hbar\omega_{ce(h)}+(n-1)[\hbar\omega_{ce}+\hbar\omega_{ch}]+\left[1-\frac{[2(n-1)]!}{2^{2(n-1)}[(n-1)!]^2}\right]E_0 \tag{10}$$

contain a contribution which is a multiple of the sum of the e and h cyclotron energies $[\hbar\omega_{ce}+\hbar\omega_{ch}]$; the last term $\approx[1-(\pi n)^{-1/2}]E_0$ for $n\gg 1$ in Eq. (10).

Evidently, the excitonic IR transitions are sensitive to the e–h Coulomb interactions. Kohn's theorem[8] is inapplicable in this situation, since the charge-to-mass ratios are different for e and h. However, as one can see from Eq. (10), the *difference*

$$E_{1s\to np^+}-E_{1s\to np^-}=\hbar\omega_{ce}-\hbar\omega_{ch} \tag{11}$$

does not depend on the e–h interactions.[4] The result (11) follows from the fact that the variables in Eq. (2) are separable in cylindrical coordinates, and it is valid not only in the limit of a high magnetic field or for a 2D system. This can likewise be attributed to the existence of an exact symmetry for excitons in a uniform field B. To show this, we introduce[9] the time-reversal operator $\hat{T}$ which operates *only* on the system under study. The field B is assumed to be an external field: The direction of **B** does not change under the operation $\hat{T}$ (the currents generating B do not change direction). In the standard manner, the coordinates do not change sign under the operation $\hat{T}$: $\hat{T}^{-1}\mathbf{r}\hat{T}=\mathbf{r}$, while the momenta and orbital angular momenta do change sign: $\hat{T}^{-1}\hat{\mathbf{p}}\hat{T}=-\hat{\mathbf{p}}$ and $\hat{T}^{-1}\hat{\mathbf{l}}\hat{T}=-\hat{\mathbf{l}}$. For the total Hamiltonian $\tilde{H}(\mathbf{K})=\hat{U}^{\dagger}(\mathbf{K})H\hat{U}(\mathbf{K})$, corresponding to the internal motion of an e–h pair (see Eq. (2)), we have

$$[\tilde{H}(\mathbf{K}),\hat{T}]=[\tilde{H}_0(\mathbf{K}),\hat{T}]=(\hbar\omega_{ch}-\hbar\omega_{ce})\hat{T}\hat{l}_z. \tag{12}$$

We shall now take into account that for excitons with $\mathbf{K}=0$ the projection l_z is a good quantum number and that $\hat{T}\Psi_{\mathbf{K}=0np^+}=\Psi_{\mathbf{K}=0np^-}$. (We note that $\hat{T}^{-1}\overline{a^{\dagger}}\hat{T}=\overline{b^{\dagger}}$, so that $\hat{T}|nm\rangle=|mn\rangle$, and the last equality is obvious in the high-field limit.) Therefore relation (11) follows from the operator algebra (12). In order for the relation (12) to hold formally it is important that the operator $\hat{T}$ is antiunitary, so that $\hat{T}^{-1}\hat{U}(\mathbf{K})\hat{T}\neq\hat{U}(\mathbf{K})$ and $[\hat{U}(\mathbf{K}),\hat{T}]\neq 0$. The analysis based on an operator algebra similar to the algebra (12) could be helpful for investigating more complicated Hamiltonians in a field B (compare with the theorem for a one-component many-electron system[9]).

5. Let us establish the characteristic features due to IR absorption by (for example, thermally excited) magnetoexcitons with $\mathbf{K}\neq 0$. We assume that the magnetic field B is high enough ($l_B\ll a_{Be(h)}$) that the mixing of different Landau levels can be taken into account by perturbation theory. The results should also be applicable qualitatively for lower fields $l_B\leqslant a_{Be(h)}$. We shall study the magnetic quantum limit $\nu_X=2\pi l_B^2 n_X\ll 1$, when magnetoexcitons fill the zeroth Landau levels; n_X is the exciton density.

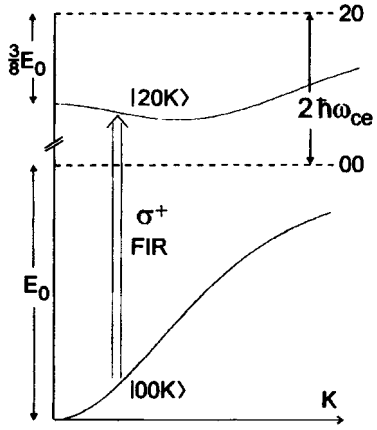


FIG. 1. Schematic illustration of the dispersion $E_{00}(K)$ and $E_{20}(K)$ of 2D magnetoexcitons $|00\mathbf{K}\rangle$ and $|20\mathbf{K}\rangle$. The vertical double arrow shows the weakly resolved σ^+ IR transition. The dashed lines mark the positions of the unoccupied Landau levels $(n_e n_h)$.

Let us consider first how the energy of a strong transition $|00\mathbf{K}\rangle \rightarrow |10\mathbf{K}\rangle$ depends on K. Assuming low temperatures $k_B T \ll E_0$, we can limit the analysis to low momenta $K l_B/\hbar \ll 1$. The dispersion relations for magnetoexcitons in this region are quadratic:[6]

$$E_{00}(K) \simeq -E_0 + K^2/2M_{00}, \quad E_{10}(K) \simeq -\frac{1}{2}E_0 + K^2/2M_{10}, \tag{13}$$

where $M_{00} = 2\hbar^2/E_0 l_B^2$ and $M_{10} = -2M_{00}$. The magnetoexciton $|10\mathbf{K}\rangle$ is characterized by a *negative* effective mass. As a result of this, the ''kinetic'' energies of the initial and final states do not compensate each other, and the transition energy

$$E_{00\rightarrow 10} = \hbar\omega_{ce} + \frac{1}{2}E_0 - \frac{K^2}{2M_{00}}\left(1 + \frac{M_{00}}{|M_{10}|}\right) \tag{14}$$

decreases with increasing momentum K. (A similar situation for the transition $|00\mathbf{K}\rangle \rightarrow |20\mathbf{K}\rangle$ is shown in Fig. 1.) Therefore it can be expected that as the temperature increases in a high magnetic field, the line due to this transition will broaden predominantly into the region of *lower* energies. Since the dispersion of 2D magnetoexcitons is due to only e–h interactions,[6] this effect is simply due to the influence of interparticle interactions on intraband excitonic IR transitions.

Let us now estimate the characteristic size of the third term in Eq. (14). Since 2D magnetoexcitons form an almost ideal gas,[10] we propose for them a Bose distribution function $f_X = (\exp[(\epsilon_K - \mu)/k_B T] - 1)^{-1}$, where $\epsilon_K = K^2/2M_{00} = E_0 K^2 l_B^2/4\hbar^2$ from Eq. (13), and the chemical potential of a 2D ideal Bose gas is given by the expression $\mu = k_B T \ln[1 - \exp(-E_0 \nu_X/2k_B T)]$. Therefore, for 2D magnetoexcitons, the particular regime which is realized is determined by the parameter $\zeta \equiv E_0 \nu_X/k_B T$. In the classical limit, $\zeta \ll 1$, we have Maxwell–Boltzmann statistics, and $\langle K^2\rangle/2M_{00} = k_B T$. In the degen-

erate quantum limit, $\zeta \gg 1$ (when the chemical potential $\mu = -k_B T e^{-\zeta/2}$ is exponentially small), we obtain $\langle K^2 \rangle / 2M_{00} = \pi^2 k_B T/3\zeta \ll k_B T$, i.e., narrowing of the absorption line occurs.

Another feature associated with IR absorption by magnetoexcitons with $\mathbf{K} \neq 0$ is due to mixing of different Landau levels. In the high magnetic field limit, when mixing is neglected, the magnetoexciton wave functions $|nm\mathbf{K}\rangle$ are given by expression (4). In the next order in the parameter $l_B/a_{Be(h)} \ll 1$, the wave functions assume the form $|\widetilde{nm}\mathbf{K}\rangle = \Sigma_{n'm'} A^{(nm)}_{n'm'} |n'm'\mathbf{K}\rangle$, where the coefficients $A^{(nm)}_{nm} = \mathcal{O}(1)$ and

$$A^{(nm)}_{n'm'} = \frac{U^{n'm'}_{nm}(\mathbf{K})}{\hbar\omega_{ce}(n-n') + \hbar\omega_{ch}(m-m')} \sim \frac{l_B}{a_{Be(h)}} \ll 1. \tag{15}$$

Here $U^{n'm'}_{nm}(\mathbf{K}) = \langle n'm'\mathbf{K}|U_{eh}|nm\mathbf{K}\rangle$ is the Coulomb matrix element between two magnetoexcitonic states. An analytical expression for $U^{n'm'}_{nm}(\mathbf{K})$ with arbitrary indices is obtained in Ref. 11 (see also Ref. 6). For magnetoexcitons with $\mathbf{K} \neq 0$ the e–h interaction mixes *all* states on different Landau levels $|nm\mathbf{K}\rangle$. This gives rise to a number of new lines in the spectra [cf. Eq. (6)]: $\langle \widetilde{nm}\mathbf{K}|\delta\hat{V}^{\pm}|\widetilde{00}\mathbf{K}\rangle \neq 0$. However, all transitions with $|n-m| \neq 1$ are found to be weak, of order $\sim (l_B/a_{Be(h)})^2$. Furthermore, for $k_B T \ll E_0$, the larger the difference $|n-m|$, the weaker the transition is. Let us consider as an example the transition $|\widetilde{00}\mathbf{K}\rangle \to |\widetilde{20}\mathbf{K}\rangle$ (see Fig. 1). We underscore that for $\mathbf{K} = 0$ this is a strictly forbidden transition $1s \to 3d^+$. The total intensity of the transition $|\widetilde{00}\mathbf{K}\rangle \to |\widetilde{20}\mathbf{K}\rangle$ (the total absorbed power is $\simeq 2\hbar\omega_{ce} R_{20}$)

$$R_{20} = \frac{2\pi}{\hbar} \sum_{\mathbf{K}} |\langle \widetilde{20}\mathbf{K}|\delta\hat{V}^{+}|\widetilde{00}\mathbf{K}\rangle|^2 f_X(K,T) \tag{16}$$

depends on the population of different $\mathbf{K}$ states; $f_X(K,T)$ is the Bose distribution function of 2D magnetoexcitons in the zeroth Landau level. In the classical ($\zeta \ll 1$) and quantum ($\zeta \gg 1$) limits we obtain

$$R_{20} = \frac{25}{32} \frac{e^2 \mathcal{F}_0^2}{\hbar} \zeta \left(\frac{k_B T}{\hbar\omega_{ce}} \right)^2 \sim T B^{-5/2}, \quad \zeta \equiv E_0 \nu_X / k_B T \ll 1, \tag{17}$$

$$R_{20} = \frac{25}{32} \frac{e^2 \mathcal{F}_0^2}{\hbar} \left[\frac{\pi^2}{3} - \zeta e^{-\zeta/2} \right] \left(\frac{k_B T}{\hbar\omega_{ce}} \right)^2 \sim T^2 B^{-2}, \quad \zeta \gg 1. \tag{18}$$

It is interesting to note that in the classical limit (17) the total intensity $R_{20} \sim n_X$, whereas in the quantum limit (18) R_{20} saturates and (to within exponential corrections) is independent of the exciton density n_X. Transitions to higher Landau levels $|\widetilde{00}\mathbf{K}\rangle \to |\widetilde{nm}\mathbf{K}\rangle$ (i.e., transitions in the σ^+ polarization with $N \equiv n-m > 1$) are suppressed even more strongly at low temperatures: For example, for $\zeta \ll 1$ their total intensity is

$$R_{nm} \sim \frac{\nu_X (k_B T)^{N-1}}{[(n+m-1)\hbar\omega_c]^2 E_0^{N-3}} \sim T^{N-1} B^{-(N+3)/2}, \tag{19}$$

where we have set as a simplification $\omega_{ce} \simeq \omega_{ch} = \omega_c$.

6. In summary, we have studied the internal magnetooptic transitions of 2D excitons. It was established that for excitons with center-of-mass momentum $\mathbf{K}=0$ the spectra contain pairs of transitions differing in energy by the difference of the cyclotron energies of an electron and hole $\hbar(\omega_{ch}-\omega_{ce})$. This result was obtained for the case of simple bands with quadratic dispersion relations. A recent experiment[12] showed that this property also holds approximately for quasi-2D excitons in a GaAs/GaAlAs quantum well with a complicated valence band. This situation will be studied theoretically in a separate publication. It was predicted that for magnetoexcitons with $\mathbf{K}\neq 0$ the spectra of strong transitions will broaden into the region of low energies with increasing temperature. It was also shown that transitions for which $\mathbf{K}\neq 0$ and which are weakly resolved are sensitive to the magnetoexciton statistics. Thus a study of transitions of this kind could be helpful in the investigation of the condensation of 2D magnetoexcitons.

I thank G. E. W. Bauer and A. Yu. Sivachenko for helpful discussions and B. D. McCombe for communicating the results of Ref. 12 prior to publication. This work is supported by grants from the Russian Fund for Fundamental Research and INTAS.

[1] M. Salib, H. A. Nickel, G. H. Herold *et al.*, Phys. Rev. Lett. **77**, 1135 (1996).
[2] J. Cerne, J. Kono, M. S. Sherwin *et al.*, Phys. Rev. Lett. **77**, 1131 (1996).
[3] A. A. Dremin, V. B. Timofeev, D. Birkedal, and J. Hvam, *Proceedings OECS-5*, Göttingen, 1997, to be published in Phys. Status Solidi A.
[4] A. B. Dzyubenko and A. L. Yablonskii, JETP Lett. **64**, 213 (1996).
[5] L. P. Gor'kov and I. E. Dzyaloshinskiĭ, Zh. Éksp. Teor. Fiz. **53**, 717 (1967) [Sov. Phys. JETP **26**, 449 (1967)].
[6] I. V. Lerner and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **78**, 1167 (1980) [Sov. Phys. JETP **51**, 588 (1980)].
[7] J. E. Avron, I. W. Herbst, and B. Simon, Ann. Phys. (N.Y.) **114**, 431 (1978).
[8] W. Kohn, Phys. Rev. **123**, 1242 (1961).
[9] A. B. Dzyubenko and A. Yu. Sivachenko Phys. Rev. B **48**, 14690 (1993).
[10] I. V. Lerner and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **80**, 1488 (1981) [Sov. Phys. JETP **53**, 763 (1981)]; A. B. Dzyubenko and Yu. E. Lozovik, Fiz. Tverd. Tela (Leningrad) **25**, 1519 (1983) [Sov. Phys. Solid State **25**, 874 (1983)]; Fiz. Tverd. Tela (Leningrad) **26**, 1540 (1984) [Sov. Phys. Solid State **26**, 938 (1984)].
[11] M. Graf, P. Vogl, and A. B. Dzyubenko, Phys. Rev. B **54**, 17003 (1996).
[12] H. A. Nickel, G. H. Herold, M. S. Salib *et al.*, *Proceedings EP2DS-12*, Tokyo, 1997, to be published in Physica B.

Translated by M. E. Alferieff

Shape of the Cu(2) NQR spectra in $YBa_2Cu_3O_7$, $TmBa_2Cu_3O_7$ and $TmBa_2Cu_4O_8$

A. V. Dooglav, A. V. Egorov, E. V. Krjukov, Yu. A. Sakhratov, and M. A. Teplov
Kazan State University, 420008 Kazan, Russia

Yu. Xu
Experimental Physik-Institut, Universität des Saarlandes Im Stadtwald, D-66123 Saarbrücken, Germany

(Submitted 30 September 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 594–598 (10 November 1997)

We present a study of shape of the Cu(2) NQR spectra in $YBa_2Cu_3O_7$, $TmBa_2Cu_3O_7$, and $TmBa_2Cu_4O_8$ compounds at temperatures of 4.2–300 K. The results of the quantitative analysis lead us to conclude that the shape of the Cu(2) NQR spectra in all the samples studied can be described in the framework of the ''motional narrowing'' model, which implies that the Cu(2) nucleus possesses two different NQR frequencies between which it can rapidly jump. The difference in frequencies seems to be related to the charge-stripe correlations in CuO_2 planes resulting in a dynamical modulation of the electric field gradients at the Cu(2) nuclei. © *1997 American Institute of Physics.*
[S0021-3640(97)00521-5]

PACS numbers: 76.60.Gv, 74.72.Bk, 74.72.Yg

In the present study an attempt was undertaken to analyze quantitatively a resonance line shape of so-called ''planar'' copper nuclei belonging to CuO_2 planes. The main problems preventing the correct studies of the planar copper NQR/NMR line shape are widely believed to arise from an enormously strong inhomogeneous broadening of spectral lines due to the large number of defects of the crystal lattice (oxygen vacancies and interstitials, twin boundaries, impurity phases, stacking faults, etc.) typical for the layered cuprates. These difficulties ultimately seemed to preclude cooperative studies of the line shape problem, since the Cu(2) NQR/NMR spectra of the same compounds prepared in different laboratories usually appeared to be very much different and ''sample-dependent.'' In order to minimize the problems of the inhomogeneous broadening due to crystal lattice defects, we have studied nominally pure stoichiometric 123 and 124 compounds. Furthermore, by measuring the Cu(2) NQR spectra (i.e., in a zero external magnetic field), the additional complications resulting from inhomogeneous broadening due to the vortex lattice in a superconducting state have been avoided. The principal finding of the paper is that the shape of the Cu(2) NQR spectra in all the samples studied can be described in a framework of the ''motional narrowing'' model,[1–3] which implies that the Cu(2) nucleus possesses two different NQR frequencies between which it can rapidly jump.

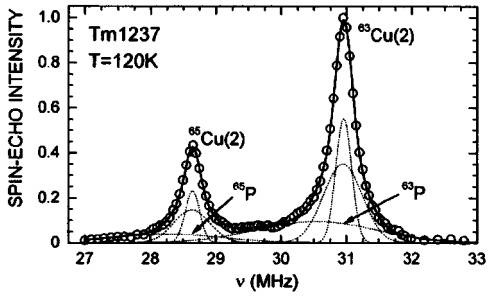


FIG. 1. The Cu(2) NQR spectrum in Tm1237 at $T=120$ K; the solid line is a best fit by six Gaussians (dotted lines; for details see Ref. 3).

All three samples studied in the present work were previously used in our NMR/NQR experiments: $YBa_2Cu_3O_7$ (Y1237),[3] $TmBa_2Cu_3O_7$ (Tm1237),[4] and $TmBa_2Cu_4O_8$ (Tm1248);[5] the critical temperatures T_c(onset) of 92.5 K, 91.5 K and 80.0 K, respectively, have been obtained in ac susceptibility versus T measurements at a frequency of 1 kHz (amplitude ≈ 1 Oe). A home-built spin-echo coherent pulsed spectrometer was used for the Cu(2) NQR spectra measurements. Both of the 1237 compounds were found to be in a slightly overdoped state. As an example, the Tm1237 spectrum is shown in Fig. 1. It is seen that, except for the relatively narrow ^{63}Cu(2) and ^{65}Cu(2) NQR lines, the spectrum of Tm1237 (like that of Y1237)[3] has a broad ''pedestal'' (**P**). It has recently been suggested[3] that the **P** spectrum arises from copper nuclei which are located in areas with a partially disordered oxygen sublattice of CuO basal planes (for example, in regions around twin boundaries) and, accordingly, with a reduced and locally inhomogeneous hole concentration in the CuO_2 planes. The spectrum of the Tm1248 sample has no such pedestal. Subtracting **P** from the observed ''raw'' spectrum, we obtain the spectrum of a ''good'' 1237 superconductor, free (or almost free) of crystal structure defects. In fact, this refinement procedure makes it possible to perform a subsequent quantitative analysis of the Cu(2) NQR line shape of a 1237 superconductor having an undistorted or slightly distorted orthorhombic structure. Fitting of the Cu(2) NQR spectra of Y1237 by using a superposition of Gaussian-type lines has shown[3] that the **P** spectrum can be satisfactorily described by a single Gaussian with a temperature-independent rms width of 1.5(1) MHz. The corresponding width of the Tm1237 pedestal turned out to be somewhat bigger, 1.9(1) MHz, whereas the relative intensities of the **P** spectra in both the Y1237 and Tm1237 samples were found to be approximately equal to 1/3.

The refined spectra of Y1237 and Tm1237 (the experimental spectra minus the ''pedestals'') and the ''raw'' spectrum of Tm1248 are shown in Fig. 2 for the ^{63}Cu isotope. A common property of all the lines in Fig. 2 is that their shape is intermediate between Gaussian and Lorentzian. Such a shape, being atypical for resonance lines in rigid-lattice solids, can be regarded as a hint that some kind of a charge motion takes place in the CuO_2 planes. Assuming that this motion is indeed present in the high-T_c cuprates, one can immediately find a qualitative explanation for two puzzles that have existed from the very beginning of the high-T_c story. First, the striking fact that the electric field gradient at the orthorhombic Cu(2) sites has axial symmetry ($\eta \approx 0$) can be

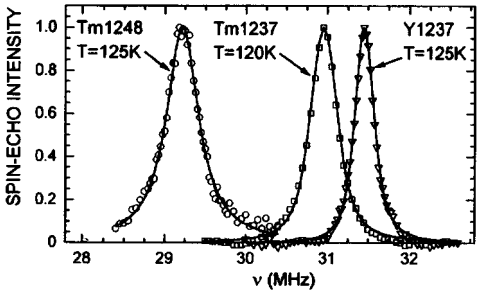


FIG. 2. The ^{63}Cu(2) NQR lines in Y1237, Tm1237 (as obtained by subtraction of the ''pedestal'' **P** from the experimental spectra; see Fig. 1 and the text for details), and Tm1248. The solid lines are a best fit by Eqs.(1) and (2) with $k=2$, $l=2$; for other parameters of the model see Fig. 3.

understood as simply the result of motional averaging. Second, the strong inhomogeneous broadening of the Cu(2) NQR lines which is observed at low temperatures can then be regarded as resulting not only from the appearance of some excess distortions of the crystal lattice but also from changes in the characteristics of the charge motion. In what follows, we try to analyze the shape of the Cu(2) NQR lines (Fig. 2) using a model with two NQR frequencies (ω_1, ω_2) between which the Cu(2) nucleus can rapidly jump. When applying the ''motional narrowing model'' we expect to find it capable of apprehending the essential difference in the Cu(2) NQR parameters at temperatures above and below T_c. For this simplest version of the model, the shape of the Cu(2) NQR spectrum can be described by the following expressions:[3]

$$S(\omega)\sim\int I(\omega,\Omega)\exp[-(\Omega-\omega_0)^2/2\sigma^2]d\Omega, \tag{1}$$

$$I(\omega,\Omega)\sim[(\omega_2-\omega_1)^2(W_2+W_1)]/[(\omega-\omega_1)^2(\omega-\omega_2)^2+(W_1(\omega-\omega_2)+W_2(\omega-\omega_1))^2]. \tag{2}$$

Here $\omega_1=\Omega+lD$, $\omega_2=\Omega-\Delta$, the frequency shifts $(+l\Delta)$ and $(-\Delta)$ are due to fluctuations of a hole density in the first and second states, respectively, $1/W_i$ is the lifetime of the ith state, $W_1=kW$, and $W_2=W$. The individual NQR line is assumed to be inhomogeneously broadened due to crystal lattice defects, so its shape is Gaussian with an rms half-width σ, and $\omega_0=\langle\Omega\rangle$ is the mean NQR frequency over the sample volume. One may regard the above model as being oversimplified, since the crystalline defects should also result in a random distribution of the other parameters, i.e., Δ, W, k and l. However, it was shown recently[3] that the Cu(2) NQR spectrum shape in the Pr-doped Y1237 compound can be fitted well by Eqs. (1) and (2) at the values $k=l=2$ corresponding to the particular conformation of charge stripes in the CuO_2 planes.[6] Therefore, we start with the same model $k=2$, $l=2$, which is expected to give some averaged values of Δ and W.

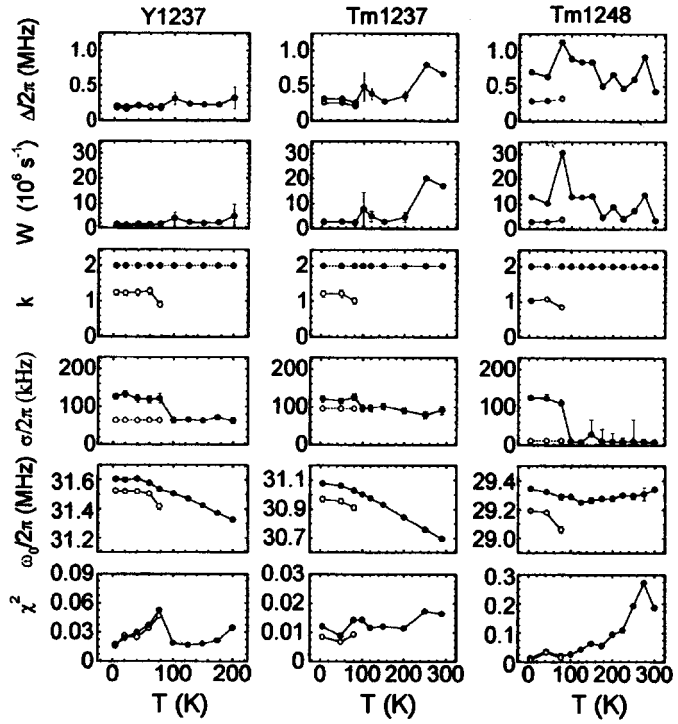


FIG. 3. The temperature dependences of the parameters $\Delta/2\pi$, W, k, $\sigma/2\pi$, and $\omega_0/2\pi$ of Eqs.(1) and (2), as obtained from fitting of the ''refined'' ^{63}Cu(2) NQR lines in Y1237 and Tm1237 and of the ''raw'' ^{63}Cu(2) NQR line in Tm1248. Filled circles correspond to a model with $k=2$, $l=2$, and variable σ; the unfilled circles correspond to a model with variable k, $l=2$, and $\sigma=$const ($\sigma_n=\sigma_s$).

The temperature dependences of the parameters $\Delta/2\pi$, W, $\sigma/2\pi$, and $\omega_0/2\pi$ for the ^{63}Cu(2) NQR lines in all three samples are shown in Fig. 3 by the filled circles, and examples of the calculated line shapes are illustrated by the solid curves in Fig. 2. For all the samples, the case of an intermediate jumping rate is realized,

$$(\omega_1-\omega_2)/W\sim 1, \tag{3}$$

and the rate itself, W, appears to be rather low: $W=2\times10^6$–2×10^7 s^{-1}. In fact, according to Eq. (3) the jumping rate is closely related to the frequency difference $\omega_1-\omega_2=3\Delta$. It is interesting to note here that the values of $(\omega_1-\omega_2)/2\pi$ appear to lie in the frequency range from 0.7 MHz (separation of the ^{63}Cu(2) NQR lines in the Pr-doped Y1237)[3] to 2.2 MHz (separation of the A and B lines in $La_{2-x}Sr_xCuO_4$ and $La_2CuO_{4+\delta}$).[7] In two of three samples studied, Y1237 and Tm1237, the difference $(\omega_1-\omega_2)$ seems to exhibit a sharp decrease at the superconducting transition. However, the most striking result is that in all the samples under study the inhomogeneous linewidth σ undergoes a sharp increase at T_c, so that one actually has two different values of σ, i.e., σ_n for $T>T_c$ and σ_s ($>\sigma_n$) for $T<T_c$, which seem to be temperature-independent. The small width σ_n in Tm1248 is close to that usually observed in the

^{63}Cu(2) NMR spectra of the crystallographically perfect materials in a high external magnetic field. For the 1237 compounds, which are known to have many structural defects, the values of σ_n appear to be rather large. As to the values of σ_s, it is found to be the same for all three samples, $\sigma_s \approx 0.8 \times 10^6$ s^{-1}. The latter fact can be regarded as a hint that the broadening of the Cu(2) NQR line at $T < T_c$ reflects some intrinsic property of high-T_c cuprates. In principle, the broadening of the copper NQR lines can originate from disordering of both the positions of the atoms and the charges of the ligands. Since the values of σ_n in 1237 compounds are much bigger than that in Tm1248, they can be naturally attributed to a disorder of the oxygen positions in the CuO basal planes. It is known[8] that even in the almost stoichiometric Y123-6.98 single crystal the chain oxygen is statically displaced in the *a* direction by 0.074(10) Å. Starting from this fact, we then arrive at the conclusion that the same type of oxygen displacements should exist in Tm1248 below T_c. In fact, the dynamic displacements of the chain oxygens in the *a* direction by 0.1 Å resulting in the formation of ferroelectric domains, were found to be present in the Y1248 compound.[9] If those displacements are indeed responsible for the broadening of the Cu(2) NQR line in Tm1248, one can conclude from the static nature of the linewidth σ_s that the oxygen motion in CuO chains slows down or even freezes at $T < T_c$. The modification of charge motion in CuO chains at $T = T_c$, if it exists, may have an indirect effect on the electronic state of CuO_2 planes via an abrupt re-distribution of holes between chains and planes.

Alternatively, the broadening of the Cu(2) NQR spectra at $T < T_c$ can be explained as arising from an abrupt re-distribution of charges in CuO_2 planes. In the particular conformation of charge stripes in Refs. 3–6, two types of Cu(2) ions are distinguished (see Fig. 2b and c in Ref. 6) — those located at the center of the stripe (type 1, the hole density on the nearest oxygen ligands is high) and those at the stripe boundaries (type 2, the hole density is low). At the optimal doping of the CuO_2 planes by holes (i.e., for close packing of the stripes in rows -2-1-2-2-1-2-) the number of centers of type 2 (n_2) appears to be twice as large as the number of centers of type 1 (n_1), in which case $k = W_1/W_2 = n_2/n_1 = 2$. Fitting of the experimental ^{63}Cu(2) NQR lines by Eqs. (1) and (2) with k and l as free parameters (along with Δ, W, σ) has shown the parameter l to be temperature-independent and close (on average) to the value of 2. When performing the subsequent fits with a constant $l = 2$ and variable k, we obtained a striking result: it turned out that the experimental data for $T < T_c$ can be well described (unfilled circles in Fig. 3) by using a constant value of $\sigma_s = \sigma_n$, the only condition necessary for this being $k \sim 1$. At $T < T_c$ the parameter χ^2 for the model $k = 1$, $l = 2$ is definitely smaller than that for $k = 2$, $l = 2$. In the framework of the model under discussion this last result can be interpreted as hinting of a modification of a stripe pattern at the superconducting transition. In particular, a stripe conformation of the -1-2-1-2-1-2- type or a checkerboard pattern can be deduced from the above condition $k = 1$. The decrease of k from 2 (above T_c) to 1 (below T_c) may actually mean that the lifetime of Cu(2) ions in the state with a high NQR frequency (corresponding to a high local density of holes at neighboring oxygen ions in the CuO_2 plane) becomes longer in the superconducting state. Moreover, the frequency difference ($\omega_1 - \omega_2$) appears to be smaller at $T < T_c$ which, perhaps, indicates that the charge-stripe modulation in the superconducting state is weaker than that in the normal state or that, in other words, charge-density waves in the CuO_2 planes coexist but compete with superconductivity.

When interpreting the experimental results on the Cu(2) NQR in the framework of the stripe model of Ref. 6 we do not rule out the possibility of other modifications for charge-density waves in the CuO_2 planes. It should be noted, however, that the above model seems to get an indirect confirmation in inelastic neutron scattering (INS) studies[10–13] of $YBa_2Cu_3O_{6.6}$. Indeed, if one takes every third hole-rich stripe in Fig. 2b of Ref. 6 away, the mean hole concentration p becomes equal to $(2/3)\times(1/6)=1/9$ per CuO_2 unit, and, according to the empirical formula $p=0.187-0.21\delta$ (Ref. 14), the oxygen index $7-\delta=6.64$ appears to be close to 6.6. The resulting stripe pattern should give rise to magnetic neutron scattering not at $\mathbf{Q}_{AF}=(1/2,1/2)$ but instead at $(1/2\pm\delta,1/2\pm\delta)$, with $\delta=1/18=0.0556$. Very recent INS experiments with $YBa_2Cu_3O_{6.6}$ single crystals have revealed such an incommensurate structure with $\delta=0.057\pm0.006$.[13]

In conclusion, analysis of the shape of the ^{63}Cu(2) NQR spectra in Y1237, Tm1237 and Tm1248 compounds performed on the basis of a simplified model of motional narrowing supports the idea that the Cu(2) nucleus possesses two different resonance frequencies between which it can jump. The plausible cause of this difference in frequencies seems to be related to charge-density waves (or dynamic charge-stripe correlations) in the CuO_2 planes, resulting in a dynamic modulation of the electric field gradients at the Cu(2) nuclei.

This work was supported in part by the Russian Scientific Council on Superconductivity, under Project 94029, and by the Russian Fund for Fundamental Research, under Project 96-02-17058a. Stimulating discussions with A. S. Borovik-Romanov, A. F. Andreev, and V. F. Gantmakher are gratefully acknowledged. The authors are also grateful to D. Brinkmann, M. Mali, and J. Roos for insightful questions and critical remarks.

[1] P. W. Anderson, J. Phys. Soc. Jpn. **9**, 316 (1954).
[2] C. P. Slichter, *Principles of Magnetic Resonance*, 3rd ed., New-York: Springer-Verlag, 1992.
[3] M. A. Teplov, Yu. A. Sakhratov, A. V. Dooglav *et al.*, JETP Lett. **65**, 821 (1997).
[4] M. A. Teplov, A. V. Dooglav, E. V. Krjukov *et al.*, JETP Lett. **82**, 370 (1996).
[5] M. A. Teplov, E. V. Krjukov, A. V. Dooglav *et al.*, JETP Lett. **63**, 227 (1996).
[6] O. N. Bakharev, M. V. Eremin, and M. A. Teplov, JETP Lett. **61**, 515 (1995).
[7] B. W. Statt, P. C. Hammel, Z. Fisk *et al.*, Phys. Rev. B **52**, 15575 (1995).
[8] P. Schweiss, W. Reichardt, M. Braden *et al.*, Phys. Rev. B **49**, 1387 (1994).
[9] T. R. Sendyka, W. Dmowski, T. Egami *et al.*, Phys. Rev. B **51**, 6745 (1995).
[10] J. M. Tranquada, P. M. Gehring, G. Shirane *et al.*, Phys. Rev. B **46**, 5561 (1992).
[11] B. J. Sternlieb, J. M. Tranquada, G. Shirane *et al.*, Phys. Rev. B **50**, 12915 (1994).
[12] J. M. Tranquada, ''Charge stripes and spin correlations in copper-oxide superconductors,'' cond-mat/9702117.
[13] P. Dai, H. A. Mook and F. Dogan, ''Incommensurate magnetic fluctuations in $YBa_2Cu_3O_{6.6}$,'' cond-mat/9707112.
[14] J. L. Tallon, C. Bernhard, H. Shaked *et al.*, Phys. Rev. B **51**, 12911 (1995).

Published in English in the original Russian journal. Edited by Steve Torstveit.

Bound states of an electron in an impurity potential on the surface of liquid helium

P. D. Grigor'ev[a)]

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia; Max-Planck-Institut für Festkorperforschung, BP166, F-38042 Grenoble, France

(Submitted 5 July 1997; resubmitted 1 October 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 599–604 (10 November 1997)

The energies and widths of the levels of an electron on impurity centers on the surface of liquid helium are calculated with allowance for the deformation of the surface. The level shift associated with the deformation effects is small and decreases very slowly with increasing level number. However, even a small shift of the energy levels relative to one another affects ripplon scattering, which makes the main contribution to the level width at low temperatures. It is predicted theoretically that this width depends very strongly on the external parameters and on the level number and that a maximum obtains at a clamping field $E_\perp = 1500$ V/cm. The width of the levels of an electron in a bound state is found to be less than for free electrons. This makes it possible to perform a beautiful spectroscopic experiment. © *1997 American Institute of Physics.* [S0021-3640(97)00621-X]

PACS numbers: 67.55.Lf, 67.57.Pq, 67.40.Yv

The behavior of electrons on the helium surface has been studied for almost 20 years. An extensive analysis of the phenomena arising here can be found in Ref. 1. In the present letter we solve the single-particle problem of an electron in an impurity potential. This is an important subject, since scattering by nonuniformities of the substrate and, especially, localization on positive ions have a strong influence on the dynamical properties of the surface electrons[2] and on the collective effects in the two-dimensional system formed by these electrons.

If a positively charged impurity is present at some distance from the helium surface, then electrons form bound states near it. The stability of such a system was investigated in Ref. 3, where the position of the impurity was determined as the result of the competition between the image force repelling the impurity from the helium–air boundary and the Coulomb force attracting the impurity to the electron on the surface. In our case the impurity is at rest at the bottom of the vessel, since it is confined by the much stronger force of the electrostatic image at the helium–substrate boundary (the permittivity of the substrate material is ordinarily greater than that of helium). Furthermore, the position of the impurities and their density can be set at a prescribed level by illuminating the substrate with a laser. In this case the question of the equilibrium position of the impurity becomes irrelevant.

Similar bound states arise at the boundary between two semiconductors, for which

Vinter[4] found by numerical methods the energy levels and the approximate form of the wave functions of the bound states for small distances of the impurity from the surface, $d<60$ Å. The situation is somewhat different in helium. First, the permittivity of helium is very close to 1 and the image forces are not so strong. Second, the electrons above helium interact with ripplons and with helium vapor, which in the present problem results in broadening of the levels. Third, there arises a static deformation of the helium surface that influences the position of the levels.

ENERGY LEVELS AND WAVE FUNCTIONS

Neglecting scattering and static deformation of the surface, which will be examined in detail below, the electron is in a potential[1]

$$V=eFz-\frac{e^2}{4z}\frac{\varepsilon-1}{\varepsilon+1}-\frac{e^2}{4(d+z)}\frac{\varepsilon_s-1}{\varepsilon_s+1}+V_0\theta(-z)-\frac{e^2Z_{\text{eff}}}{\sqrt{(d+z)^2+x^2+y^2}},$$

where the permittivity of helium is $\varepsilon=1.045$. The effective charge of the impurity is $Z_{\text{eff}}=(2/\varepsilon+1)(2\varepsilon/\varepsilon_s+\varepsilon)Z$. Even without an external field ($F=0$) an electron is clamped very strongly to the surface (the average distance from the surface is $\langle z\rangle<100$ Å), so that $\langle z\rangle/d\ll 1$. Therefore the last term, which corresponds to the impurity, can be assumed to be independent of the coordinate z and to depend only on its average value $\tilde{z}$. This approximation also works well because of the fact that the impurity potential is a small correction to the potential along the z axis (but by no means in a plane parallel to the helium surface). Now the variables separate; the wave function has the form $\psi(x,y,z)=\zeta(z)f(x,y)$, where $\zeta(z)$ and $f(x,y)$ satisfy the equations

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dz^2}+eFz-\frac{e^2}{4z}\frac{\varepsilon-1}{\varepsilon+1}-\frac{e^2}{4d}\frac{\varepsilon_s-1}{\varepsilon_s+1}+V_0\theta(-z)-E_n\right)\zeta(z)=0, \tag{1}$$

$$\left(-\frac{\hbar^2}{2m}\left(\frac{d^2}{dx^2}+\frac{d^2}{dy^2}\right)-\frac{e^2Z_{\text{eff}}}{\sqrt{d^2+x^2+y^2}}-E_{kl}\right)f(x,y)=0. \tag{2}$$

In the last equation we introduced the notation $d=d+\tilde{z}$, where $\tilde{z}\simeq\langle z\rangle=\int\zeta^*(z)\zeta(z)z\,dz$. In the absence of a field one has $\langle z\rangle=a=3\hbar^2/2m\alpha\approx 100$ Å and $\alpha=(\varepsilon-1)e^2/4(\varepsilon+1)$. For very strong fields in the bottom subband $\langle z\rangle=b\approx 1.6(\hbar^2/2meF)^{1/3}$. On this basis, $\tilde{z}=(a^{-2}+b^{-2})^{-1/2}$ gives fairly good accuracy (the error $\leqslant 1\%$ of the value of d).

The splitting between the subbands (between energy levels along the z axis) is about 30 K, and at low temperatures an electron ''freezes'' in the bottom subband. For this reason, the two-dimensional behavior of an electron is determined by Eq. (2), which looks simple but cannot be solved exactly. For the lower levels, the region of localization $\langle x^2\rangle\sim\langle y^2\rangle\sim 200$ Å and an expansion can be made in powers of the parameter $\langle x^2+y^2\rangle/d^2$. Then the potential assumes the form

$$V=\frac{Z_{\text{eff}}e^2}{d}\left(-1+\frac{r^2}{2d^2}-\frac{3}{8}\left(\frac{r^2}{d^2}\right)^2+\frac{5}{16}\left(\frac{r^2}{d^2}\right)^3-\ldots\right).$$

In zeroth-order perturbation theory, we have a two-dimensional oscillator with $\omega=\sqrt{Z_{\text{eff}}e^2/d^3m}\approx(4\ \text{K})\sqrt{Z_{\text{eff}}}$.

As we can see, the corrections of higher orders will be very important in determining the width of the levels. We shall employ perturbation theory to calculate these corrections. A potential of the form r^{2n} is not diagonal in the basis of wave functions

$$f_{kl}(x,y)=\left(\frac{m\omega}{\pi\hbar}\right)^{1/2}\frac{1}{\sqrt{2^{k+l}k!l!}}e^{-(x^2+y^2)/2a^2}H_k\left(\frac{x}{a}\right)H_l\left(\frac{y}{a}\right).$$

For this reason, it is more convenient to classify the states with respect to the projection m of the angular momentum. Then the wave functions can be expressed in terms of the confluent hypergeometric function

$$f_{nm}(r,\phi)=C_{nm}e^{-x/2}x^{m/2}F(-n,m+1,x)e^{im\phi},$$

where the radial quantum number $n=0,1,\ldots,$ and $x=r^2\sqrt{(Z_{\text{eff}}e^2m_e/d^3\hbar^2)}$. The first-order correction to the energy levels is given by the integral

$$E_{kl}^{(1)}=V_{(nm)(nm)}^{(1)}=\int f_{nm}^2(r,\phi)\left(-\frac{3}{8}\frac{Z_{\text{eff}}e^2}{d^2}\right)\left(\frac{r}{d}\right)^4 r\,dr\,d\phi.$$

Integrals of this type are easy to calculate [Ref. 5, Appendix f]. The final answer is

$$\begin{aligned}E_{nm}^{(1)}=&-\frac{3}{8}\frac{\hbar^2(m+2)(m+1)}{m_ed^2}\\&\times\left[1+\sum_{s=0}^{n-1}\frac{n(n-1)\ldots(n-s)(-3-s)(-2-s)\ldots(-2+s)}{[(s+1)!]^2(m+1)(m+2)\ldots(m+s)}\right]\\&\times\left[1+\sum_{s=0}^{n-1}\frac{n(n-1)\ldots(n-s)(-1-s)(-s)\ldots(s)}{[(s+1)!]^2(m+1)(m+2)\ldots(m+s)}\right]^{-1}.\end{aligned}\tag{3}$$

For the second level, this formula gives the corrections to the energy

$$E_{10}^{(1)}=-\frac{3}{4}\frac{7\hbar^2}{m_ed^2}\quad\text{and}\quad E_{02}^{(1)}=-\frac{3}{4}\frac{6\hbar^2}{m_ed^2}.$$

Perturbation theory works if $\hbar(n^2+m^2+1)/2e\sqrt{mdZ_{\text{eff}}}\ll1$, which for $d=500$ Å holds fairly well for the six lowest levels.

DEFORMATION EFFECTS

Under the action of the clamping field an electron exerts a pressure on the helium surface, deforming it. This static, craterlike deformation of the surface acts back on the electron and localizes it in the region[1,6] $L=4\pi\alpha\hbar^2/me^2E_\perp^2$, where $E_\perp=E_0+e/d^2$ is the sum of the external field and the field of the impurity. The energy benefit $W=-(e^2E_\perp^2/4\pi\alpha)\ln(1/\kappa L)$ accompanying the formation of the crater is of the same order of magnitude as the correction $E^{(1)}$ in first-order perturbation theory in $(r/d)^2$. In

addition, W is virtually independent of the electron level number. Therefore the static deformation of the surface may be regarded as a perturbation and investigated in first order only.

The energy correction due to the formation of the crater is $W=\int d^2r(\alpha/2)\,[(\nabla\xi)^2+\kappa^2\xi^2]-eE_\perp|\Psi|^2\xi$. We need not solve the self-consistent problem of finding $\xi(r)$ and $|\Psi(r)|^2$ as was done in Refs. 1 and 6, since in first order the electron wave function $\Psi(r)$ does not depend on the surface deformation $\xi(r)$ but only on the number of the level in the impurity potential. The function $\xi(r)$ that minimizes W is given by the equation $\alpha\xi''(r)+(\alpha/r)\,\xi'-\rho g\xi=eE_\perp|\Psi(r)|^2$. The solution of this equation is

$$\xi(r)=\frac{1}{\alpha}\int_0^\infty G(\omega)J_0(\omega r)\omega d\omega, \tag{4}$$

where

$$G(\omega)=\frac{P(\omega)}{\omega^2+\kappa^2},\quad P(\omega)=\int_0^\infty eE_\perp|\Psi(r)|^2J_0(\omega r)rdr. \tag{5}$$

For the ground state of an electron with wave function $|\Psi(r)|^2=(1/\pi a^2)\exp(-r^2/a^2)$

$$G(\omega)=\frac{eE_\perp}{2\pi(\omega^2+\kappa^2)}\exp\left(-\frac{a^2\omega^2}{4}\right)\Rightarrow\xi(0)\approx\frac{eE_\perp}{4\pi\alpha}\ln\frac{4}{\gamma\kappa^2a^2}$$

where $\gamma=1.78$.[6] Since $\ln(\kappa^2a^2)\gg1$, the quantity $\xi''(0)=2eE_\perp/\pi\alpha a^2\ll\xi(0)/2a^2\Rightarrow\xi(r)$ starts to change substantially on scales larger than the electron localization region and then falls monotonically to zero over distances $1/\kappa=0.05$ cm: $\xi(r)|_{r\to\infty}\approx eE_\perp K_0(\kappa r)/2\pi\alpha$. The shift of the lower level is $W_{00}\approx eE_\perp\xi(0)\approx0.7$ K with $E_\perp=3000$ V/cm. This is a large correction, but it is still much less than the splitting between the levels in the zeroth approximation, so that perturbation theory remains applicable. Let us now see how the deformation correction depends on the level number.

It follows from Eq. (4) that $G(\omega)$ and $\xi(r)$ depend linearly on $|\Psi(r)|^2$. Therefore the value of $\xi(0)$ for different levels can be obtained by differentiating expression (5) with respect to the parameter $1/a^2$. For the level $n_r=0$, $m=1$

$$|\Psi(r)|^2_{01}=\frac{1}{\pi a^4}r^2e^{-r^2/a^2}=-\frac{1}{a^4}\frac{\partial}{\partial(1/a^2)}(a^2|\Psi|^2_{00})$$

$$\Rightarrow G_{01}=\frac{eE_\perp}{2\pi(\omega^2+\kappa^2)}\frac{-1}{a^4}\frac{\partial}{\partial(1/a^2)}\left[a^2\exp\left(-\frac{a^2\omega^2}{4}\right)\right]$$

$$=\frac{eE_\perp}{2\pi(\omega^2+\kappa^2)}\left(1-\frac{a^2\omega^2}{4}\right)\exp\left(-\frac{a^2\omega^2}{4}\right)\Rightarrow\xi_{01}=\xi_{00}-\frac{eE_\perp}{4\pi\alpha}.$$

The difference of the shifts of the first and zeroth levels is $\ln(4/\gamma\kappa^2a^2)\approx22$ times less than the shift of each level, but even a very small shift of the levels relative to one

another can affect their width, which as we shall see below is very sensitive to the splitting between the sublevels of the split level. Let us calculate the deformation correction for different sublevels of the second level:

$$G_{02}=\frac{eE_\perp}{2\pi(\omega^2+\kappa^2)}\left(1-\frac{a^2\omega^2}{2}+\frac{a^4\omega^4}{32}\right)\exp\left(-\frac{a^2\omega^2}{4}\right)\Rightarrow\xi_{02}=\xi_{00}-\frac{3}{2}\frac{eE_\perp}{4\pi\alpha},$$

$$G_{10}=\frac{eE_\perp}{2\pi(\omega^2+\kappa^2)}\left(1-\frac{a^2\omega^2}{2}+\frac{a^4\omega^4}{16}\right)\exp\left(-\frac{a^2\omega^2}{4}\right)\Rightarrow\xi_{10}=\xi_{00}-\frac{eE_\perp}{4\pi\alpha}.$$

The energy correction is $W_{02}=W_{00}+3(eE_\perp)^2/2(4\pi\alpha)$ and $W_{10}=W_{00}+(eE_\perp)^2/4\pi\alpha$. Deformation effects increase the splitting between the sublevels of the second level, decreasing its ripplon width appreciably. The calculations performed above have confirmed that the deformation correction to the energy decreases very slowly with increasing level number.

LEVEL WIDTH

Scattering by atoms in the helium vapor and by ripplons make the main contribution to the level width. In the case of free electrons the width of intersubband transitions which is due to scattering by atoms was found both theoretically and experimentally in Ref. 7. For bound electrons, the arguments remain the same and the answer differs only by a coefficient whose value is close to 1. Scattering by the vapor gives the width $\hbar\Delta\omega=m_eN_GU_G^2/2\hbar^2a$, where U_G is related with the electron scattering cross section of an atom by the formula $A=m^2U_G^2/\pi\hbar^4\approx5\times10^{-16}$ cm^2, the localization region $a=\sqrt{\hbar/m\omega}$, and the vapor density decreases very rapidly with decreasing temperature, $N_G=(Mk_BT/2\pi\hbar^2)^{3/2}\exp(-Q/k_BT)$, where $Q=7.17$ K. At $T=1$ K one has $N_G=1.5\times10^{18}$ cm^{-3} and $\Gamma\approx0.01$ K.

For free electrons the ripplon width obtained by numerical methods for an intersubband transition by T. Ando[8] decreases with temperature more slowly than does the width due to scattering by the atoms. For this reason ripplon scattering dominates at temperatures $T<0.8$ K.

Let us examine the width of discrete levels of an electron in an impurity potential. Here there are substantial differences from the case of free electrons, so that we shall perform the calculation from the very beginning. According to Ref. 1, the electron-ripplon interaction operator for transitions within the same subband has the following form in an approximation linear in the vertical displacement ξ of the surface: $\hat{V}^R=\int d^2r\psi^+(r)\psi(r)\Sigma_q\xi_qe^{i\mathbf{q}\cdot\mathbf{r}}V(q)$, where $\xi_q=Q(q)(b_q+b^+_{-q})$, $Q(q)=(\hbar q\ \tanh(qd)/2\rho\omega_q)^{1/2}$, and

$$V(q)=\int\frac{(\varepsilon-1)e^2q}{4(\varepsilon+1)z}\left(\frac{1}{qz}-K_1(qz)\right)\zeta^2(z)dz+eE_\perp.$$

The first term in $V(q)$ is due to the image-force distortions and the second is due to the clamping field. These contributions are equal for $E_\perp=300$ V/cm. Therefore we replace the first term by $eE_{\text{eff}}=300$ V/cm and introduce $F=E_\perp+E_{\text{eff}}$.

To find the width we shall calculate the probability of a transition of an electron into other levels and sum these probabilities:

$$\Gamma_n^R = 2\pi \sum_m \int \frac{d^2q}{(2\pi)^2} \langle (b_q + b_{-q}^+)(b_{-q} + b_q^+) \rangle Q^2(q) |V_{mn}^R|^2 \delta(E_m + \omega(q) - E_n).$$

The matrix element $V_{mn}(q) = \int \psi_m(r) \psi_n(r) e^{i\mathbf{q}\cdot\mathbf{r}} d^2r V(q)$ decreases very rapidly as $q \to \infty$. If the unperturbed oscillator eigenfunctions are taken for $\psi(r)$, then $V_{mn}(q) \sim e^{-q^2a^2/4}$. Analysis of the exact solution of the Schrödinger equation does not change this estimate. Thus only the interaction with the long-wavelength ripplons, whose energy is not much greater than $\hbar\omega(q = 2/a) \approx 0.02$ K but equals the energy splitting between the levels, need be taken into account. If we actually did have a harmonic oscillator, then there would be no such levels and the ripplon width would be negligibly small. This is not the case because the degenerate levels of the oscillator are split, and the width depends strongly on the magnitude of this splitting. We note that the ground level was nondegenerate and therefore did not split, and the first level was degenerate only with respect to the projection of the angular momentum and also did not split in a potential of the form r^{2k}. Let us calculate the ripplon width of the second level. The electron wave functions are

$$\psi_{02} = \frac{1}{\sqrt{2\pi a^2}} e^{-r^2/2a^2} \left(\frac{r}{a}\right)^2 e^{im\phi}, \quad \psi_{10} = \frac{1}{\sqrt{\pi a^2}} e^{-r^2/2a^2} \left(1 - \frac{r^2}{a^2}\right).$$

The matrix element is

$$|V_{(10)(02)}^R|^2 = (eF)^2 \exp\left(-\frac{q^2a^2}{2}\right) \frac{(qa)^4}{8} \left(1 - \frac{q^2a^2}{8}\right)^2$$

and the width of the second level is

$$\Gamma_2^R = \frac{(eF)^2}{\alpha} \frac{2n_q + 1}{24} \exp\left(-\frac{q_0^2a^2}{2}\right) (q_0a)^4 \left(1 - \frac{q_0^2a^2}{8}\right)^2,$$

where $n_q = 1/(e^{\hbar\omega_0/kT} - 1)$, and q_0 is determined by the relation

$$\hbar\omega(q_0) = \hbar \sqrt{\frac{\alpha}{\rho}} q_0^{3/2} = E_{10}^{(1)} + W_{10} - E_{02}^{(1)} - W_{02} = \frac{3}{4} \frac{\hbar^2}{m_e d^2} + \frac{(eE_\perp)^2}{8\pi\alpha},$$

whence

$$qa = \left(\frac{1}{\hbar}\sqrt{\frac{\rho}{\alpha}}\right)^{2/3} \left(\frac{\hbar^2 d^3}{mZe^2}\right)^{1/4} \left[\frac{3\hbar^2}{4md^2} + \frac{(eE_\perp)^2}{8\pi\alpha}\right]^{2/3}.$$

One can see that the width depends very strongly on qa, i.e., on the depth d, on the substrate permittivity ε_s, and on the level number. For $d = 1000$ Å, $F = 3000$ V/cm, $Z_{\text{eff}} = 1$, and $T = 0.5$ K we obtain $qa = 5.2$ and $\Gamma_2^R = 0.003$ K. If d is halved, then $qa = 7$ and $\Gamma_2^R \sim 10^{-5}$ K. The width of the second level is a nonmonotonic function of the clamping field. For $E_\perp \approx 1500$ V/cm it has a flat maximum which arises as a result of the competition between the factor $(E_\perp + E_{\text{eff}})^2$ and the exponential dependence of the width

on $q_0 a$, where the clamping field enters on account of the deformation corrections. For $d = 1000$ Å one has $\Gamma \approx 0.005$ K at the point of the maximum. The surface deformation strongly influences the width of the second level. This is also due to the fact that the splitting between its sublevels is small. For the third level the initial splitting is 11 times larger, and for this reason the width of this level should be very small and be more independent of deformation effects.

At low temperatures the contribution of scattering by the atoms is exponentially small, and the interaction with ripplons remains the only cause of the level broadening. The strong dependence of the ripplon width on the external parameters and on the level number is a distinguishing feature of an electron in a bound state. The smallness of the level width for $T < 0.6$ K makes it possible to perform a beautiful spectroscopic experiment, similar to that of Ref. 7 but for bound states.

I thank A. M. Dyugaev and V. B. Shikin for a helpful discussion.

[a)]e-mail: pashag@itp.ac.ru

[1] V. B. Shikin and Yu.P. Monarkha, *Two-Dimensional Charged Systems in Helium* [in Russian], Nauka, Moscow, 1989.
[2] Yu. Z. Kovdrya, F. F. Mende, and B. A. Nikolaenko, Fiz. Nizk. Temp. **10**, 1129 (1984) [Sov. J. Low Temp. Phys. **10**, 589 (1984)].
[3] Yu. P. Monarkha and Yu. Z. Kovdrya, Fiz. Nizk. Temp. **8**(2), 215 (1982) [Sov. J. Low Temp. Phys. **8**, 107 (1982)].
[4] B. Vinter, Phys. Rev. B **26**, 6808 (1982).
[5] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed., Pergamon Press, New York, 1977 [cited Russian original, 4th ed., Nauka, Moscow, 1989].
[6] V. B. Shiskin and Yu. P. Monarkha, Zh. Eksp. Teor. Fiz. **65**, 751 (1973) [Sov. Phys. JETP **38**, 373 (1974)].
[7] C. C. Grimes, T. R. Brown, Michael L. Burns, and C. L. Zipfel, Phys. Rev. B **13**, 140 (1976).
[8] Ts. Ando, J. Phys. Soc. Jpn. **44**, 765 (1978).

Translated by M. E. Alferieff

Aharonov–Bohm oscillations in a ring with a quantum well

I. A. Ryzhkin
Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow District, Russia

(Submitted 23 September 1997; resubmitted 1 October 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 605–610 (10 November 1997)

Aharonov–Bohm oscillations in a ring with a quantum well are investigated in the ballistic regime. It is shown that when trajectories with multiple circuits around the ring are taken into account, the maxima in the conductivity correspond to resonance levels of an isolated ring. The results obtained are in qualitative agreement with the experiment performed by Yakoby, Heiblum, Mahalu, and Shtrikman [Phys. Rev. Lett. **74**, 4047 (1995)]: Although the scattering phase of an electron scattered by a quantum well changes by π on passage through each resonance, the Aharonov–Bohm curves for the centers of neighboring resonances are identical. In the simplified interpretation employed by Yakoby *et al.* the latter result looks like an identical scattering phase in neighboring resonances. © *1997 American Institute of Physics.*
[S0021-3640(97)00721-4]

PACS numbers: 72.15.Rn, 73.90.+f

The development of nanotechnology in the last 10 years has led to the creation of electronic structures several tens of nanometers in size (quantum wires, contacts, wells, and their combinations). On account of their smallness and low impurity density, the transport in such structures at low temperature is ballistic: Along their entire propagation path electrons do not undergo any scattering by phonons or a random static impurity potential. The only form of scattering is scattering by elements of the structure itself. Under such conditions electron propagation is a quantum-limit process and is described best by the quantum theory of scattering and not by the classical Boltzmann equation. Under quantum scattering conditions the behavior of a scattered electron is characterized not only by the scattering probability but also by the scattering phase. For this reason the conductivity of topologically nontrivial structures can depend directly on the phase. This is a new circumstance compared to classical transport, and it can serve as a basis for the development of fundamentally new electronic devices with unusual characteristics.

The electrical properties of one such structure were investigated in detail in Ref. 1 and shown schematically in Fig. 1: a ring with current contacts and a quantum well in one of the arms. The ring was placed in a magnetic field and thus carried a magnetic flux Φ. A voltage U was applied to one of the electrodes creating the quantum well, thereby changing the depth of the well. The conductance of the structure as a function of the flux Φ manifested Aharonov–Bohm oscillations with period equal to the magnetic flux quantum $\Phi_0 = hc/e$, indicating transport coherence, i.e., transport was sensitive to the phase

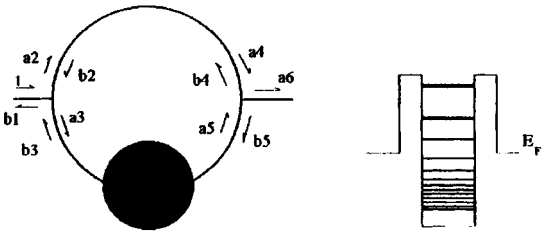


FIG. 1. Aharonov–Bohm ring with a quantum well in the bottom arm (on the left-hand side of the figure). The magnetic flux is directed upwards out of the plane of the figure. The arrows show the incident and reflected waves. A quantum well with resonance levels is shown schematically on the right-hand side.

acquired in a magnetic field. The conductance as a function of U in the presence of zero magnetic flux had the form of narrow resonance peaks at voltages which the authors associated with matching of the resonance levels of the well with the Fermi level of the system. The most interesting and contradictory results were obtained in an investigation of the scattering phases of the quantum well. There the Aharonov–Bohm oscillations were investigated at fixed voltages, and the scattering phases were determined from the shifts of the oscillations relative to the origin of coordinates. It was found that when the energy is scanned through each resonance the scattering phase changes by π, as should happen according to the Breit–Wigner formula.[2] However, the phases in successive resonances turned out to be the same, which is obviously at variance with simple models of a quantum well and resonance levels (the phases should differ by π). This contradiction, formulated on the basis of a simple and graphic theory, appears to be completely unresolvable if the electron–electron interaction is neglected. It is probably for this reason that the latest attempts to resolve it[3–5] are all based on models of interacting electrons.

The objective of the present letter is to make a theoretical analysis of the conductance of an Aharonov–Bohm ring with a quantum well in one arm of the ring and to investigate the dependence of the conductance on the magnetic field and voltage applied to the well. In contrast to Refs. 3–5, we employ the same model of noninteracting electrons as that in Ref. 1, but we do not confine ourselves to summing only the two simplest trajectories connecting the contacts 1 and 2. Actually, our method of solution is equivalent to taking into account nonrectilinear trajectories, including those with different numbers of circuits around the ring. When they are taken into account, propagation along the two arms is no longer independent and actually signifies that a necessary condition for a resonance in the conductance is that the Fermi level coincides with the resonance level of the entire ring (isolated from the current contacts) and not with the resonance level of the quantum well. This circumstance radically changes the outcome and gives a simple and natural explanation for all observed results.

To find the conductance, we note first that it can be expressed, by means of the Landauer formula,[6] in terms of the transmission coefficient, the finding of which is a standard quantum-mechanical problem. We represent the electron wave function in each segment in the form

$$\psi_i(x)=\frac{a_i}{\sqrt{k}}\exp(ikx)+\frac{b_i}{\sqrt{k}}\exp(-ikx), \tag{1}$$

where a_i and b_i are the amplitudes of the waves shown in the figure by arrows, x is the one-dimensional coordinate along a segment (we assume that the system is one-dimensional or has only one channel), and k is the wave number. The wave functions in each arm are matched by means of unitary scattering matrices expressing the amplitudes of the scattered waves in terms of the amplitudes of the incident waves:

$$\begin{Bmatrix} b_2 \\ a_4 \end{Bmatrix}=\begin{Bmatrix} 0 & t_0' \\ t_0 & 0 \end{Bmatrix}\begin{Bmatrix} a_2 \\ b_4 \end{Bmatrix}, \quad \begin{Bmatrix} b_3 \\ a_5 \end{Bmatrix}=\begin{Bmatrix} r & t' \\ t & r' \end{Bmatrix}\begin{Bmatrix} a_3 \\ b_5 \end{Bmatrix}, \tag{2}$$

where r and r' are the reflection coefficients and t and t' the transmission coefficients of the lower arm, treated as a single scatterer; similarly, t_0 and t_0' are the transmission coefficients for the upper arm. The upper arm is assumed to be free and the reflection coefficients for it equal zero. In the absence of a magnetic flux one has $t_0=t_0'=\exp(ikl)$, where l is the length of the semicircle. When a magnetic flux is switched on, r and r' do not change but the transmission coefficients do change: $t_0\to t_0\exp(i\alpha)$, $t_0'\to t_0'\exp(-i\alpha)$, $t\to t\exp(-i\alpha)$, and $t'\to t'\exp(i\alpha)$, $\alpha=(\Phi/\Phi_0)\pi$. The location where the current-conducting contacts connect with the ring are also described by scattering matrices (3×3 matrices).[7] Unitarity (flux conservation) and symmetry (reversibility in time), the symmetry between the two arms of the ring, and the additional requirement that the matrix elements be real allow the scattering matrices of the units to be parametrized by a single real parameter:[8]

$$\begin{Bmatrix} b_1 \\ a_2 \\ a_3 \end{Bmatrix}=\begin{Bmatrix} -a-b & \sqrt{\varepsilon} & \sqrt{\varepsilon} \\ \sqrt{\varepsilon} & a & b \\ \sqrt{\varepsilon} & b & a \end{Bmatrix}\begin{Bmatrix} 1 \\ b_2 \\ b_3 \end{Bmatrix},$$

$$\begin{Bmatrix} a_6 \\ b_5 \\ b_4 \end{Bmatrix}=\begin{Bmatrix} -a-b & \sqrt{\varepsilon} & \sqrt{\varepsilon} \\ \sqrt{\varepsilon} & a & b \\ \sqrt{\varepsilon} & b & a \end{Bmatrix}\begin{Bmatrix} 0 \\ a_5 \\ a_4 \end{Bmatrix}, \tag{3}$$

where $a,b=(\sqrt{1-2\varepsilon}\mp 1)/2$ and ε varies in the range $[0,1/2]$ and characterizes the coupling of the ring with the contacts. The value $\varepsilon=0$ corresponds to an isolated ring and $\varepsilon=1/2$ corresponds to the maximum coupling of the ring with the contacts. We underscore that the units are treated as point objects. For this reason, they do not enclose a finite magnetic flux, and the constants a, b, and ϵ do not depend on the magnetic field. We also note that an ideal connection (in the sense that the fluxes are conserved and the wave function is continuous) corresponds to a matrix (3) with $\varepsilon=4/9$ which is symmetric with respect to all three arms. The system of equations (2) and (3) makes it possible to find directly the transmission amplitude ($=a_6$). Omitting long algebraic calculations, we present the final expression for the transmission amplitude F_{12} for small ε, which is the most interesting and, from the experimental standpoint, most likely case:

$$F_{12}=\varepsilon\left[\frac{t_0(1-t^2+r+r'+rr')\exp(i\alpha)+t(1-t_0^2)\exp(-i\alpha)}{D_0+\varepsilon D_1}\right], \tag{4}$$

where D_0 and D_1 are determined by the equations

$$D_0=(1-tt_0\ \exp(-2i\alpha))(1-tt_0\ \exp(2i\alpha))-t_0^2rr'. \tag{5}$$

$$D_1=2tt_0\ \cos(2\alpha)+(r+r')(1+t_0^2)/2+2t_0^2(rr'-t^2). \tag{6}$$

Using the Landauer formula, the equations (4)–(6), and the unitarity of the scattering matrices, the conductance G_{12} of the ring can be represented in the form (here and below the conductance is given in units of $2e^2/h$)

$$G_{12}=\varepsilon^2\frac{[r\ \cos(\beta)-\sin(\delta)]^2+t^2\ \sin^2(\varphi)-2t\ \sin(\varphi)[r\ \cos(\beta)-\sin(\varphi)]\cos(2\alpha)}{[\cos(\delta+\varphi)-t\ \cos(2\alpha)]^2+\varepsilon^2[(r\ \cos(\beta)-\sin(\delta))\ \cos(\varphi)-\sin(\varphi)\cos(\delta)]^2}. \tag{7}$$

Here and below r and t are the moduli of the reflection and transmission coefficients for scattering by a quantum well, the value of the parameter $\beta=[\arg(r')-\arg(r)]/2$ depends on the model of the quantum well and in the general case is nonzero, δ is the scattering phase of the quantum well, and $\varphi=kl$ is the phase acquired during free propagation along the top arm. It is important to underscore that besides the symmetry dictated by time reversal $G_{12}(\Phi)=G_{21}(-\Phi)$, the conductance exhibits an additional symmetry $G_{12}(\Phi)=G_{12}(-\Phi)$.

It is evident from Eq. (7) that for small ε the dependence of the conductance on the parameters of the problem is of a resonance form. Maxima occur when $D_0=0$ (the first term in the denominator in Eq. (7) vanishes). It is easy to show that this condition is identical to the condition for quantization of the levels of an isolated ring. The height of a maximum is independent of ε, whereas the width is directly proportional to ε. Therefore the conductance resonances correspond to the energy levels of the ring and not the quantum well by itself. A striking illustration of this assertion is the possibility of a conductance resonance even in regions far from the well resonances, where $t=0$ (this requires that $\cos(\delta+\varphi)=0$). The conductance at resonances of this type does not depend on the magnetic flux. This is entirely understandable, since in the case when propagation occurs only along one arm there is no interference mechanism there. This case probably does not correspond to the experiment of Ref. 1.

Strong modulation of the conductance by a magnetic flux (pronounced Aharonov–Bohm oscillations) exists only for t comparable to 1, i.e., when the Fermi level is close to a resonance level of the well (but does not necessarily coincide with it). However, it would be incorrect to think that a neighborhood of each resonance level of the well necessarily results in a conductance resonance for a fixed magnetic flux. Indeed, suppose that δ_r and t_r from a neighborhood of a resonance of the well give a conductance resonance, i.e., they are solutions of the equation $\cos(\delta+\varphi)-t\cos(2\alpha)=0$. Then the corresponding point of the next well resonance $(\delta_r+\pi)$, t_r is not a solution of this equation with the same magnetic flux. It is easy to see, however, that it becomes a solution when the magnetic flux changes by $\delta\Phi=\Phi_0/2$ and thereby gives a resonance with a different value of the magnetic field. We underscore that the heights and widths of

the conductance resonances at these successive points can be completely different (for some parameters of the problem some of them can be observed and others cannot).

The magnetic flux dependence of the conductance in Eq. (7) appears at two locations: in the numerator and in denominator. The term with $\cos(2\alpha)$ in the numerator is virtually identical to the result of the simplified theory[1] and originates from a simple addition of the amplitudes of the waves propagating along the two arms. This dependence is weak, of the order of ε^2. The denominator in Eq. (7) (or equivalently Eq. (4)) is actually obtained by adding all trajectories with different numbers of circuits around the ring and leads to a more complicated and stronger dependence on account of the sharp resonance form of Eq. (7).

Next let us investigate the dependence of the conductance on the direction and form of the Aharonov–Bohm oscillations near the resonance levels of the quantum well. The moduli of the reflection and transmission coefficients and the scattering phase can be expressed as

$$r=\frac{|\eta|}{\sqrt{\eta^2+1}},\quad t=\frac{1}{\sqrt{\eta^2+1}},\quad \delta=\pm\frac{\pi}{2}-\arctan(\eta), \tag{8}$$

where $\eta=(E_F-E_0)/\gamma$ is the dimensionless deviation of the resonance level E_0 of the well from the Fermi level, γ is the width of the resonance level, and the $\pm$ are chosen in order to permit comparison of two successive resonances differing in phase by π. Substituting expression (8) into Eq. (7) gives for the conductance

$$G_{12}=\varepsilon^2\frac{[|\eta|\cos(\beta)\mp 1]^2+\sin^2(\varphi)-2[|\eta|\cos(\beta)\mp 1]\sin(\varphi)\cos(2\alpha)}{[\pm\eta\cos(\varphi)\pm\sin(\varphi)+\cos(2\alpha)]^2+\varepsilon^2[(|\eta|\cos(\beta)\mp 1)\cos(\varphi)\pm\eta\sin(\varphi)]^2}. \tag{9}$$

The voltage and magnetic flux dependences of the conductance are strongly determined by the value of the phase φ, which in turn depends on the dimensions of the ring and the Fermi energy. For $\cos(\varphi)=0$ the sharp η dependence observed experimentally does not occur. To come closer to experiment, let us examine the directly opposite case $\cos(\varphi)=1$. We obtain for the conductance

$$\varepsilon^2 G_{12}=\frac{[|\eta|\cos(\beta)\mp 1]^2}{[\pm\eta+\cos(2\alpha)]^2+\varepsilon^2[|\eta|\cos(\beta)\mp 1]^2}. \tag{10}$$

It is evident from this formula that a conductance resonance exists only in a quite close neighborhood of a well resonance, $|\eta|\leqslant 1$. As the energy is scanned within one well resonance in the region $-1\leqslant\eta\leqslant 1$, the value of 2α changes by π, which corresponds to the Breit–Wigner formula and the experimental result. At the same time, for $|\eta|>1$ the Aharonov–Bohm oscillations become increasingly shallow. However, if the form of the oscillations directly near the center of the resonance ($\eta\ll 1$) is investigated, then the result does not depend on which sign, upper or lower, is chosen. This means that the form of the oscillations is identical at the centers of successive well resonances (the oscillations appear to be in phase). This qualitative analysis is also confirmed by a numerical analysis of Eqs. (7), (9), and (10) for different values of the parameters β, ε, and φ.

In closing, let us list the basic results of this work. First, conductance resonances correspond to the energy levels of the isolated ring and not of the well. Second, the resonances of an isolated well should be separated into two types: even and odd (with scattering phase equal to 0 and π, respectively). If the even well resonances lead to conductance resonances with an integral magnetic flux ($\Phi = n\Phi_0$), then the odd resonances lead to conductance resonances with a half-integral flux ($\Phi = (n + 1/2)\Phi_0$). For definite parameters of the problem, the latter (or former) could be much weaker or narrower, i.e., unobservable experimentally. This is one of the possible resolutions of the paradox formulated in Ref. 1. Finally, a systematic calculation of the conductance taking into account trajectories with different numbers of circuits around the ring leads to a more complicated dependence of the conductance on the magnetic flux and the voltage than the simplified theory of Ref. 1. This dependence in itself looks like a ''change in scattering phase by π'' near each resonance and ''identical phase'' at the centers of successive resonances. The latter is the most likely explanation of the experimental results obtained in Ref. 1.

I thank V. I. Marchenko, V. M. Édel'shteĭn, and G. B. Lesovik for valuable discussions and interest in this work. This work was supported by the Russian Fund for Fundamental Research under Grant No. 96-02-19568.

[1] A. Yakoby, M. Heiblum, D. Mahalu, and H. Shtrikman, Phys. Rev. Lett. **74**, 4047 (1995).
[2] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Nonrelativistic Theory*, 2nd ed., Pergamon Press, New York, 1965 [Russian original, Nauka, Moscow, 1963].
[3] A. L. Yeyati and M. Buttiker, Phys. Rev. B **52**, R14360 (1995).
[4] G. Hackenbroich and H. A. Weidenmuller, Phys. Rev. Lett. **76**, 110 (1996).
[5] C. Bruder, R. Fazio and H. Schoeller, Phys. Rev. Lett. **76**, 114 (1996).
[6] R. Landauer, Philos. Mag. **21**, 863 (1970).
[7] B. Shapiro, Phys. Rev. Lett. **50**, 747 (1983).
[8] M. Buttiker, Y. Imry and M. Ya. Azbel, Phys. Rev. A **30**, 1982 (1984).

Translated by M. E. Alferieff

Change in the nature of the Ni diffusion mechanism on the Si(111) surface with adsorption of Co atoms

A. E. Dolbak, B. Z. Ol'shanetskiĭ,[a)] S. A. Tiĭs, and R. A. Zhachuk
Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 6 October 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 611–614 (10 November 1997)

The diffusion of Ni on a Si (111) surface is investigated by LEED and Auger electron spectroscopy. It is found that, in contrast to the process on the initially clean Si (111)–7×7 surface, on Si (111) surfaces with submonolayer Co coverages the nature of the Ni transport mechanism changes at a temperature of about 750 °C, and the Ni surface diffusion coefficients increase sharply below this temperature. © *1997 American Institute of Physics.* [S0021-3640(97)00821-9]

PACS numbers: 68.35.Fx, 82.80.Pv, 61.14.Hg

Research on surface diffusion is important for surface physics and its applications in technology. It is known that surface diffusion can depend on surface orientation and structure, the densities of atomic steps, the chemical nature and density of adsorbed atoms, external fields, and so on (see, for example, Refs. 1–4).

Using LEED and Auger electron spectroscopy (AES), we have observed an abrupt change in the nature of the mechanism of surface diffusion of Ni, with a sharp increase in the Ni surface diffusion coefficients on the Si(111) surface at temperatures below 70 °C upon the adsorption of submonolayer quantities of Co atoms as compared with Ni diffusion on the initially clean Si(111)–7×7 surface. In a previous work[5] we investigated in detail Ni diffusion on a clean Si surface. In the present work we performed only control experiments on a clean silicon surface.

The temperature dependences of the nickel surface diffusion coefficients $D(T)$ were calculated from the concentration distributions $C_{\mathrm{Ni}}(x)$ (x is the distance from the edge of the strip) obtained on a Si(111) surface as a result of the diffusion of Ni from a nickel strip deposited on the surface, as the sample is annealed for a time t. The distributions $C_{\mathrm{Ni}}(x)$ were measured by the AES method at room temperature. The transport of Ni on a clean silicon surface was observed at temperatures above 700 °C. At this temperature, after the sample was annealed for 60 min, we could not record the concentration distributions $C_{\mathrm{Ni}}(x)$ because the sensitivities and spatial resolution of the AES and LEED methods are too low. It follows from our experiments that transport of Ni atoms along clean Si surfaces occurs by means of diffusion of the atoms through the bulk followed by segregation on the surface as a result of a decrease in the solubility of Ni in Si with decreasing temperature of the sample.[5] The conclusion that Ni diffuses through the bulk of the Si is based on the fact that the concentration distributions of Ni on the Si surface appeared only after the sample cooled, while no Ni Auger signal was detected from the

surface when Ni diffused during annealing of the sample, though during this time interval Si Auger peaks as well as Ni Auger peaks from the surface of the deposited strip could be observed. The conclusion that diffusion occurs through the bulk also agrees with the fact that the nickel diffusion coefficients do not depend on the orientation and structure of the silicon surface;[5] such a dependence should be present if the nickel atoms diffuse along the surface. Moreover, this conclusion was indirectly confirmed by the fact that the values of the Ni diffusion coefficients calculated from our experimental results are close to the corresponding values presented in Refs. 6–8 for the diffusion coefficients of Ni diffusing along interstices in Si. Our conclusion about the mechanism of the transport of nickel on a clean silicon surface was confirmed in Ref. 9.

Nickel nonetheless still does diffuse on clean silicon surfaces. This is seen as the formation of nickel-induced ordered surface structures and formation of epitaxial islands of nickel disilicide $NiSi_2$. But the Ni surface diffusion coefficients are several orders of magnitude lower than the diffusion coefficients along interstices, and surface diffusion makes a negligibly small contribution to Ni transport on clean Si surfaces. This result is remarkable in that it does not fit into the existing picture of surface diffusion, according to which the migration rates of atoms along a surface are higher and the activation energy of the surface diffusion is lower than the corresponding values in the bulk of a crystal.

To investigate the effect of Co adsorption on Ni surface diffusion, a submonolayer Co coating (in what follows, a monolayer is denoted as ML) was deposited on a clean Si(111) surface in an ultrahigh vacuum at room temperature. As we showed in Ref. 10, heating of a Si(111) surface on which Co is adsorbed results in the formation of a number of Co-induced surface structures which depend on the coverage and the conditions of heat treatment. These are surface structures Si(111)–7×7–Co, Si(111)–1×1–Co, Si(111)–$\sqrt{7}\times\sqrt{7}$–Co, and Si(111)–$\sqrt{13}\times\sqrt{13}$–Co. They are observed at room temperature and are formed in a definite range of Co concentrations and annealing temperatures. If the Co concentration exceeds the amount required for a given surface structure to form, the excess Co atoms combine with silicon to form epitaxial islands of cobalt disilicide $CoSi_2$ (Ref. 10). The LEED patterns from the Si(111)–7×7–Co surface are similar to those observed from the clean Si(111)–7×7 surface. This is due to the fact that Co atoms adsorbed on a Si(111)–7×7–Co surface are incorporated in the epitaxial $CoSi_2$ islands, which occupy a small area and a large part of the silicon surface remains clean. The curves of the Ni surface diffusion coefficients were measured on samples with surface structures 7×7–Co(0.4 and 0.8 ML Co), 1×1–Co (0.6 ML Co), and $\sqrt{7}\times\sqrt{7}$–Co (0.2 ML Co). Before the Ni diffusion experiments were performed, samples with $\sqrt{7}\times\sqrt{7}$–Co and 1×1–Co surface phases and Co coatings of 0.2 and 0.6 ML, respectively, were annealed at temperatures up to 600 °C for 180 min. In the process, we did not observe any changes in the LEED patterns. This attests to the stability of the surfaces of the samples during the anneals. The values obtained for Ni diffusion coefficients on the $\sqrt{7}\times\sqrt{7}$–Co surface exhibited poor reproducibility, and we do not present them in this letter.

The concentration distributions $C_{\mathrm{Ni}}(x)$ measured on clean Si(111) surfaces and on Si(111) surfaces with adsorbed cobalt were found to satisfy the equation

$$C(x)=C_0\ \mathrm{erfc}(x/2\sqrt{Dt}). \tag{1}$$

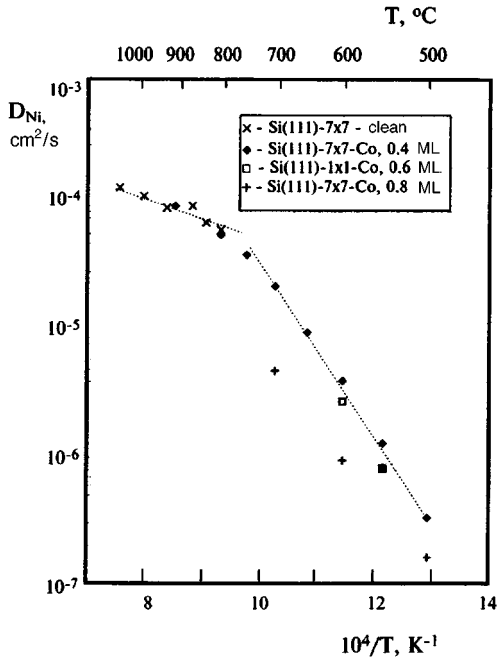


FIG. 1. Temperature dependence of the Ni surface diffusion coefficients on a clean Si(111) surface and a Si(111) surface with adsorbed Co.

Similar distributions are characteristic for one-dimensional surface diffusion from a source of constant intensity.[11]

The temperature dependences of the Ni diffusion coefficients on clean Si(111) surfaces and Si(111) surfaces with adsorbed Co are displayed in Fig. 1. The diffusion coefficients are determined with an accuracy of $\pm 30\%$ from the distributions $C_{\mathrm{Ni}}(x)$ measured at the given temperature.

The curves obtained for a clean silicon surface at temperatures above 700 °C are identical to those obtained for a silicon surface with adsorbed Co. However, below 750 °C a sharp change occurs in the mechanism of nickel transport along a silicon surface. Nickel diffusion is not detected on a clean surface below 750 °C, whereas concentration distributions $C_{\mathrm{Ni}}(x)$ are detected on a surface with adsorbed Co all the way up to a temperature of 500 °C. At about 750 °C the slope of the temperature dependence of the diffusion coefficients changes. The activation energy of Ni surface diffusion calculated for the Si(111)$-7\times7-$Co surface (see Fig. 1) equals 1.3 eV, and the temperature dependence of the diffusion coefficients can be expressed as $D=1\times10^{2}\exp(-1.3/kT)$ cm^2/s. The corresponding expression for the Ni diffusion coefficients measured on the clean Si(111) surface has the form $D=2.4\times10^{-3}\exp(-0.32/kT)$ cm^2/s.[5]

As follows from our experiments, the Ni diffusion mechanism on a Si(111) surface with adsorbed Co is different from the mechanism observed on a clean silicon surface. When submonolayer quantities of Co are present on the surface, an Auger signal from Ni

is observed in the process of diffusion of Ni at temperatures of 500–700 °C, and the nickel concentration distribution observed on the surface of a heated sample at the completion of annealing is identical to that measured at room temperature. Therefore Ni transport occurs in this case by means of diffusion along the surface, in contrast to the clean silicon surface, where, as we have said, nickel diffuses through the bulk of the Si and is then segregated at the surface.

It was natural to expect that the Ni surface diffusion coefficients would depend on the structure of the silicon surface and the concentration of adsorbed cobalt. One can see from the figure that the Ni surface diffusion coefficients on a Si(111) surface with different surface structures and adsorbed Co concentrations in the temperature range 500–700 °C differ somewhat: The Ni diffusion coefficients are lower on surfaces with a high Co concentration.

In summary, we have observed that adsorption of submonolayer concentrations of Co atoms results in an abrupt change in the nature of the mechanism of Ni diffusion on a Si(111) surface at a temperature of about 750 °C. The values of the Ni diffusion coefficients on a Si(111) surface with adsorbed cobalt are much higher than the corresponding values on an initially clean silicon surface in the temperature range 500–700 °C. However, at present we cannot explain the mechanism responsible for the increase in the Ni surface diffusion coefficients on a Si(111) surface containing cobalt-induced surface phases in the temperature range 500–700 °C. Especially surprising is the effect of Co atoms on the nickel diffusion mechanism on the Si(111)–7×7–Co surface, where the Co atoms are incorporated into irregularly arranged epitaxial islands of the disilicide $CoSi_2$, which occupy a small fraction of the surface, while a large part of the silicon surface remains clean and even retains the 7×7 structure of the clean surface.

This work was supported by the Russian Fund for Fundamental Research (Grant 95-02-05336) and the Russian State Program ''Surface Atomic Structures'' (95-1.17).

[a)]e-mail: olshan@isp.nsc.ru

[1] A. G. Naumovets and Yu. S. Vedula, Surf. Sci. Rep. **4**, 365 (1984).
[2] Ya. E. Geguzin and Yu. S. Kaganovskiĭ, *Diffusion Processes on Crystal Surfaces* [in Russian], Énergiya, Moscow, 1984.
[3] Yu. S. Vedula, A. T. Loburets, and A. G. Naumovets, Zh. Éksp. Teor. Fiz. **77,** 773 (1979) [Sov. Phys. JETP **50**, 391 (1979)].
[4] Yu. L. Gavrilyuk and V. G. Lifshtiz, Poverkhnost' **4**, 143 (1983).
[5] A. E. Dolbak, B. Z. Olshanetsky, S. I. Stenin *et al.*, Surf. Sci. **218**, 37 (1989).
[6] J. H. Aalberts and M. L. Verheijke, Appl. Phys. Lett. **1**, 19 (1962).
[7] M. Yoshida and K. Furusho, Jpn. J. Appl. Phys. **3**, 531 (1964).
[8] Eicke R. Weber, Appl. Phys. A **30,** 1 (1983).
[9] M. Y. Lee and P. A. Bennett, Phys. Rev. Lett. **75**, 4460 (1995).
[10] A. E. Dolbak, B. Z. Olshanetsky, and S. A. Teys, Surf. Sci. **373**, 43 (1997).
[11] B. I. Boltaks, *Diffusion in Semiconductors*, Academic Press, New York, 1963 [Russian original, Nauka, Moscow, 1961].

Translated by M. E. Alferieff

Magnetic structure of iron-doped UPd_2Ge_2

A. M. Balagurov and V. V. Sikolenko
Laboratory of Neutron Physics, Joint Institute of Nuclear Research, 141980 Dubna, Russia

I. S. Lyubutin[a)]
Institute of Crystallography, Russian Academy of Sciences, 117333 Moscow, Russia

G. André and F. Bourée
Laboratoire Leon Brillouin, CEA-CNRS, CE-Saclay, 91191 Gif-sur-Yvette, France

H. M. Duh
Department of Physics, National Kaohsiung Normal University, Kaohsiung, Taiwan 802

(Submitted 9 October 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 615–619 (10 November 1997)

Neutron diffraction investigations have shown that a 2% substitution Fe atoms for Pd radically alters the magnetic structure of UPd_2Ge_2. If the magnetic structure in the undoped compound at $T<50$ K consists of a longitudinal spin density wave (LSDW) with ''square'' modulation, then in polycrystalline $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ a ''simple'' antiferromagnetic (AF) phase is observed below 65 K and a sinusoidally modulated LSDW–AF phase is observed between 65 K and the Néel temperature $T_N=135$ K. In the interval $65<T<135$ K the magnetic cell is incommensurate with the crystal cell, with the exception of the point $T=93$ K, where the wave vector of the magnetic structure passes through a ''commensurate'' value equal to 0.75. Below T_N the magnetic moments of the uranium atoms are always parallel to the tetragonal axis c of the unit cell. © *1997 American Institute of Physics.*
[S0021-3640(97)00921-3]

PACS numbers: 75.50.Ee, 75.30.Fv

The first neutron diffraction investigations of UPd_2Ge_2 (Ref. 1) revealed the presence below $T_N=140$ K of a magnetic structure of the antiferromagnetic (AF) type with a modulated longitudinal spin density wave (LSDW). Recent magnetic measurements[2–5] have indicated the existence of at least two additional magnetic transitions — near 87 K and 50 K — and the possible coexistence of several magnetic phases at low temperatures — AF, ferromagnetic (FM), metamagnetic (MM), and spin glass (SG). This has stimulated repeated neutron diffraction investigations of UPd_2Ge_2 (Ref. 6) with much better resolution and accuracy than in the early work.[1] It was shown that at temperatures $1.4\leqslant T\leqslant 50$ K the magnetic structure of UPd_2Ge_2 is described by a commensurate LSDW with ''square'' modulation and with a magnetic wave vector $\mathbf{k}=(0,0,3/4)$. At temperatures $95\leqslant T\leqslant T_N$ ($T_N=135$ K) the magnetic structure consists of an incommensurate LSDW with sinusoidal modulation. In the interval $50\leqslant T\leqslant 95$ K the structure consists of a mixture of these two phases and gradually transforms from a square-

modulated into a sinusoidally modulated structure; a first-order phase transition from the commensurate into an incommensurate phase, expressed as a sharp change in k_z, occurs near 80 K. Only the uranium atoms carry a magnetic moment, and the direction of the moments are parallel to the c axis of the unit cell.

The system UPd_2Ge_2–UFe_2Ge_2 forms a continuous series of solid solutions with a structure of the $ThCr_2Si_2$ type.[3,4] It is curious that the final member of this series, UFe_2Ge_2, is nonmagnetic and is a Pauli paramagnet.[7,4] Measurements of the magnetic susceptibility and magnetization in the system $U(Pd_{1-x}Fe_x)_2Ge_2$ have shown[3,4] that even a very low iron concentration radically changes the magnetic state of the initial compound UPd_2Ge_2. Specifically, 2% Fe is a critical concentration at which a number of magnetic properties change radically.

We performed neutron diffraction magnetic and structural investigations of the system $U(Pd_{1-x}Fe_x)_2Ge_2$ for different Fe and Pd concentrations. The present letter reports the results for the critical concentration $x=0.02$.

Polycrystalline samples of pure UPd_2Ge_2 and of the doped compound $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ were prepared by melting stoichiometric quantities of the constituent materials in an argon arc furnace. The details of the sample preparation method are described in Ref. 3. The magnetic susceptibility and magnetization measurements were performed on a SQUID magnetometer in the temperature interval 5–300 K and in external magnetic fields H up to 5.5 T.

Neutron diffraction data for refinement of the structure were obtained for several temperatures on a high-resolution Fourier diffractometer in the IBR-2 pulsed reactor in Dubna. A G4.1 diffractometer in the Orphee reactor in Saclay, France, with a neutron wavelength of 2.426 Å was used for the neutron magnetic investigations. All neutron diffraction patterns obtained on the G4.1 diffractometer were recorded in a regime with temperature increasing from 1.4 K up to 142 K. The data were analyzed by the Rietveld method using the MRIA and FullProf codes.[8,9]

Figure 1 displays the temperature dependences of the magnetic susceptibility χ for pure and iron-doped samples. The inset in Fig. 1 also shows curves of the inverse susceptibility $1/\chi=f(T)$. Besides the magnetic anomaly at T_N, two additional anomalies are observed in both samples. In UPd_2Ge_2 the susceptibility passes through a maximum at $T_{m1}=87$ K and $T_{m2}\approx 50$ K. In $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ the maximum of the susceptibility at T_{m1} shifts to 74 K and becomes very intense, while the maximum at $T_{m2}\approx 50$ K is strongly suppressed.

High-resolution neutron diffraction patterns confirm that $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ possesses a simple body-centered tetragonal structure of the $ThCr_2Si_2$ type with space group I4/mmm and atoms in the following crystallographic positions: U in 2(a): 0, 0, 0; Pd (Fe) in 4(d): 0, 1/2, 1/4; Ge in 4(e): 0, 0, z. The z coordinate of Ge and the thermal factors for all the atoms were refined, and no substantial anomalies were observed in their temperature dependence.

Only the coordinate z(Ge) and the lattice parameters were refined with the nuclear diffraction peaks obtained on the G4.1 diffractometer. It was found that $z(\mathrm{Ge})=0.3812(2)$ at 142 K, and this parameter remains constant all the way down to 1.4 K. A small

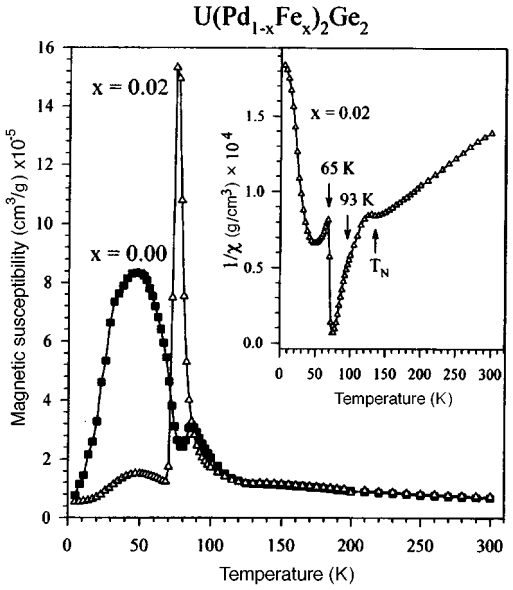


FIG. 1. Temperature dependence of the magnetic susceptibility in the field $H=2$ kOe for pure UPd_2Ge_2 and iron-doped $U(Pd_{0.98}Fe_{0.02})_2Ge_2$. Inset: Inverse susceptibility of $U(Pd_{0.98}Fe_{0.02})_2Ge_2$.

anomaly was observed in the temperature dependences of the lattice parameters a and c near 65 K. This anomaly was not observed[6] in pure UPd_2Ge_2.

It follows from the neutron data that at $T<65$ K the compound $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ possesses a simple AF structure with wave vector $\mathbf{k}=\mathbf{c}^*$. The magnetic moments of uranium are parallel to the c axis, and the magnetic unit cell coincides with the nuclear cell. The only difference between the cells is that body-centered symmetry does not hold for the magnetic cell (the magnetic moments are oppositely oriented at the points (0, 0, 0) and (1/2, 1/2, 1/2) and, in consequence, the magnetic peaks with $h+k+1\neq 2n$ are visible in the low-temperature diffraction spectra.

Figure 2 shows the temperature variations of the intensity of the magnetic peaks (100) and (101^-), referring to the commensurate and incommensurate phases, respectively. The transition at $T=65$ K to the incommensurate structure is formally manifested as a sharp decrease in the z component of the wave vector $\mathbf{k}$ from 1 to 0.76 (in units of c^*) (see Fig. 3). At temperatures $T>65$ K the magnetic structure of $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ consists of an incommensurate longitudinal spin density wave (LSDW) with sinusoidal modulation. The phases AF and LSDW coexist in a very narrow temperature interval near 65 K.

We observed another curious effect: Above 65 K, as the wave vector decreases slowly (Fig. 3), its length passes through the ''commensurate'' value of 3/4 at 93 K. Therefore $k_z>3/4$ at $65<T<93$ K and $k_z<3/4$ at $93\text{ K}<T<T_N$.

The temperature dependence of the magnetic moment of uranium in the $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ sample is shown in Fig. 4. As the temperature increases, the magni-

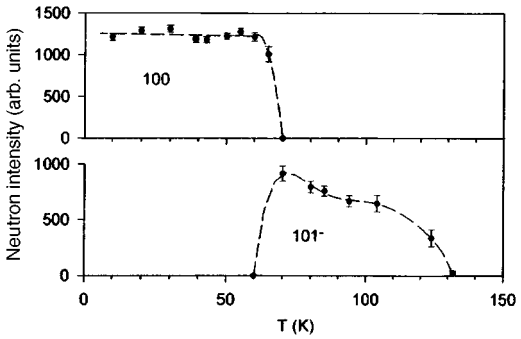


FIG. 2. Temperature dependence of the intensity of the magnetic Bragg peaks 100 and 101^{-}. The peak 100 vanishes at $T \approx 65$ K, while the peak 101^{-} appears at the same temperature.

tude of the moment remains almost constant in the region $0 < T < 65$ K. Then it decreases abruptly at 65 K, and on the way to the Néel temperature $T_N = 135$ K it passes through an additional anomaly near 93 K. This temperature is also the point where k_z becomes equal to the ''commensurate'' value of 3/4.

The neutron diffraction experiments show that even very light doping with iron changes the magnetic structure of UPd_2Ge_2 radically. After doping with only 2% Fe the low-temperature LSDW phase of pure UPd_2Ge_2 with ''square'' modulation ($k_z = 3/4$) changes into a simple AF ($k_z = 1$) structure. As the temperature increases, the magnetic transition from the commensurate to the incommensurate phase in UPd_2Ge_2 is accompanied by a sharp decrease in the number k_z (Ref. 6), while in the doped sample the wave number k_z varies continuously in the entire temperature range from 65 K up to 135 K. One gets the impression that in $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ the wave number, in going from the

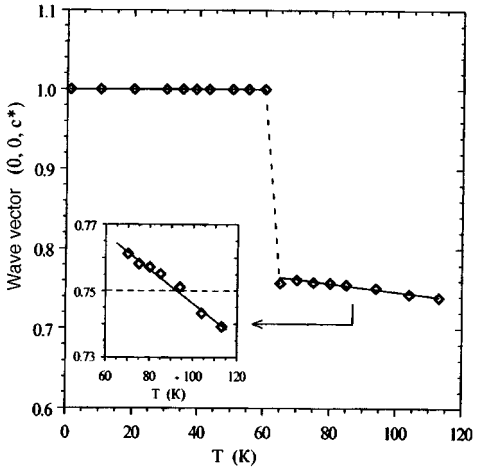


FIG. 3. Temperature dependence of the component k_z of the magnetic wave vector **k** for $U(Pd_{0.98}Fe_{0.02})_2Ge_2$. Inset: Transition of the wave vector k_z through the ''commensurate'' value ($k_z = 0.75$) at $T = 93$ K.

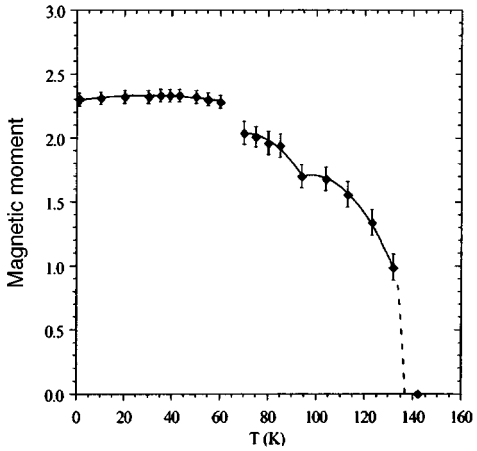


FIG. 4. Temperature dependence of the magnetic moment of uranium (in units of μ_B) in $U(Pd_{0.98}Fe_{0.02})_2Ge_2$.

value $k_z > 3/4$ to $k_z < 3/4$ at 93 K, does not at all ''feel'' this temperature as a phase-transition point. However, the anomaly in the magnitude of the magnetic moment of uranium at $T \approx 93$ K (Fig. 4) confirms the existence of a magnetic transition. It should be noted that the sharp change in the wave vector **k** and magnitude of the magnetic moment of uranium at the transition point $T = 65$ K correlate with the anomalies of the lattice parameters a and c at this temperature.

The magnetic transition temperatures 65 and 93 K found for $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ in the diffraction experiment are different from the positions of the maxima of the magnetic susceptibility (50 and 74 K; Fig. 1), but they correspond to the onset of the rise (67 K) and the termination of the drop (94 K) of the susceptibility in the peak at T_{m1}. According to Fig. 1, $\chi(T)$ increases rapidly at the point $T = 67$ K, which according to the neutron data is the temperature of the transition AF $\Rightarrow$ LSDW. One can also see in the inset in Fig. 1 that the value of $1/\chi$ increases in the interval 74–135 K, with a change in slope occurring exactly at the point $T = 93$ K. Therefore, like the magnetic moment of the uranium, the magnetic susceptibility is also sensitive to the transition $k_z > 3/4 \Rightarrow k_z < 3/4$.

It is known that several crystals with the $ThCr_2Si_2$ structure possess a simple AF structure at low temperatures, for example, UNi_2Si_2 at $53 < T < 103$ K,[10] UPd_2Si_2 at $T < 108$ K,[11] and UNi_2Ge_2 at $T < 80$ K,[12] but this structure has not been observed[6] in pure UPd_2Ge_2. We have now found that the low-temperature phase of iron-doped UPd_2Ge_2 possesses a simple AF structure.

To relate the neutron diffraction data with the results of magnetic measurements it is necessary to assume[3,4] that the compound $U(Pd_{0.98}Fe_{0.02})_2Ge_2$ below T_N is in a multiphase state. The coexistence of different phases and the competition between them result in frustrations of the magnetic order, which results in the appearance of SG and FM phases in small volumes of the sample. The iron impurity intensifies the frustration, changes the magnetic anisotropy of the material, and strongly modifies the magnetic properties of UPd_2Ge_2. Apparently, the development of the magnetic states in

$U(Pd_{0.98}Fe_{0.02})_2Ge_2$ with temperature is determined by the temperature dependence of the anisotropy energy.[3] The neutron-diffraction experiments did not reveal SG or FM phases. To check for their presence we plan to investigate these compounds by the method of μSR spectroscopy.

This work was supported by the Russian Fund for Fundamental Research and the INTAS-RFFI Fund.

[a)]e-mail: root@magnet.crystal.msk.su

---

[1] H. Ptasiewicz-Bak, J. Leciejewicz, and A. Zygmunt, J. Phys. F **11**, 1225 (1981).
[2] C. Tien, C. S. Wur, H. M. Duh *et al.*, Solid State Commun. **89**, 171 (1994).
[3] H. M. Duh, I. S. Lyubutin, and K. D. Lain, Phys. Rev. B **52**, 4294 (1995).
[4] H. M. Duh, I. S. Lyubutin, I. M. Jiang *et al.*, J. Magn. Magn. Mater. **153**, 86 (1996).
[5] H. M. Duh, I. S. Lyubutin, C. S. Wur *et al.*, J. Phys.: Condens. Matter **7**, 2165 (1995).
[6] G. Andre, F. Bouree, A. Oles *et al.*, Solid State Commun. **97**, 923 (1996).
[7] A. J. Dirkmaat, T. Endstra, E. A. Knetsch *et al.*, J. Magn. Magn. Mater. **84**, 143 (1990).
[8] V. B. Zlokazov and V. V. Chernyshev, J. Appl. Crystallogr. **25**, 591 (1992).
[9] J. Rodrigues-Carvajal, Physica B **192**, 55 (1993).
[10] H. Lin, L. Rebelsky, M. F. Collins *et al.*, Phys. Rev. B **43**, 13232 (1991).
[11] B. Shemirani, H. Lin, M. F. Collins *et al.*, Phys. Rev. B **47**, 8672 (1993).
[12] Z. Zolnierec and J. Mulak, J. Magn. Magn. Mater. **140–144**, 1393 (1995).

Translated by M. E. Alferieff

Zero curvature representation for classical lattice sine-Gordon model via quantum *R* matrix

A. V. Zabrodin

Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia; Joint Institute of Chemical Physics, 117334 Moscow, Russia

(Submitted 1 October 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **66**, No. 9, 620–625 (10 November 1997)

Local M operators for the classical sine-Gordon model in discrete space–time are constructed by convolution of the quantum trigonometric 4×4 R matrix with certain vectors in its ''quantum'' space. Components of the vectors are τ functions of the model. This construction generalizes the known representation of continuous-time M operators in terms of the classical r matrix. © *1997 American Institute of Physics.* [S0021-3640(97)01021-9]

PACS numbers: 03.50.Kk, 11.10.Lm

1. Soliton equations are integrable Hamiltonian systems,[1] with Poisson brackets for Lax matrices having a unified form in terms of the (classical) r matrix. An alternative approach[1,2] consists in representing soliton equations as 2D zero curvature (ZC) conditions for a pair of matrices called L and M operators depending on a spectral parameter. Although this method avoids any reference to the Hamiltonian aspects, the r matrix arises here, too, as a machine to produce M operators from L operators. Let us recall how it works.

Let $\mathcal{L}_l(z)$ be a classical ultralocal 2×2 L operator on 1D lattice with the periodic boundary condition $\mathcal{L}_{l+N}(z)=\mathcal{L}_l(z)$; z is the spectral parameter. The monodromy matrix is $\mathcal{T}_l(z)=\mathcal{L}_{l+N-1}(z)\ldots\mathcal{L}_{l+1}(z)\mathcal{L}_l(z)$. Hamiltonians of commuting flows are obtained by expanding $\log T(z)$ in z, where $T(z)=\mathrm{Tr}\,\mathcal{T}_l(z)$ does not depend on l due to the periodic boundary condition. All these flows admit a ZC representation. The generating function of corresponding M operators is[3,1]

$$M_l(z;w)=T^{-1}(w)\,\mathrm{Tr}_1[r(z/w)(\mathcal{T}_l(w)\otimes I)], \tag{1}$$

where $r(z)$ is the r matrix (of size 4×4) acting in the tensor product of two 2-dimensional spaces, Tr_1 means the trace in the first space, and I is the identity matrix.

A way to construct local M operators from (1) is well known.[4,1,5] Suppose there exists a z_0 such that $\det \mathcal{L}_l(z_0)=0$ for any l, so $\mathcal{L}_l(z_0)$ is a projector:

$$\mathcal{L}_l(z_0)=\frac{|\alpha_l\rangle\langle\beta_l|}{\lambda_l},\quad |\alpha\rangle=\begin{pmatrix}\alpha^{(1)}\\ \alpha^{(2)}\end{pmatrix},\quad \langle\beta|=(\beta^{(1)},\beta^{(2)}). \tag{2}$$

Here λ_l is a scalar normalization factor. Then $M_l(z;z_0)$ is a local quantity:

$$M_l(z)\equiv M_l(z;z_0)=\frac{\langle\beta_l|r(z/z_0)|\alpha_{l-1}\rangle}{\langle\beta_l|\alpha_{l-1}\rangle}. \tag{3}$$

The scalar product is taken in the first space only, so the result is a 2×2 matrix. It obeys the ZC condition $\partial_t\mathcal{L}_l(z)=M_{l+1}(z)\mathcal{L}_l(z)-\mathcal{L}_l(z)M_l(z)$ with the spectral parameter.

The goal of this work is to extend Eq. (3) to M operators for *discrete time flows* in Hirota's 2D partial difference equations.[6–8] We follow Refs. 9 and 10, treating the discrete equations as members of the same infinite hierarchy as the continuous ones.

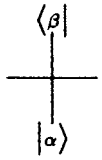
Let us outline the results. In the discrete case the function $r(z)$ in Eq. (3) is replaced by the quantum R matrix. Specifically, the following representation of discrete M operators $\mathcal{M}_l(z)$ holds:

$$\mathcal{M}_l(z)=\frac{\langle\beta_l|R(z/z_0)|\check{\beta}_{l-1}\rangle}{\langle\beta_l|\alpha_{l-1}\rangle}\quad |\check{\beta}_l\rangle\equiv\sigma_1|\beta_l\rangle \tag{4}$$

(hereafter σ_i are Pauli matrices). On the right-hand side, $R(z)$ is a *quantum* 4×4 R matrix to be specified below, with the ''quantum'' parameter related to the time lattice spacing. A similar formula for the L operator itself is valid with another quantum R matrix $R^{(-)}(z)$:

$$\mathcal{L}_l(z)=\frac{\langle\beta_l|R^{(-)}(z/z_0)|\alpha_l\rangle}{\langle\beta_l|\alpha_{l-1}\rangle}. \tag{5}$$

The vectors $|\alpha_l\rangle$ and $|\beta_l\rangle$ are *the same* as in Eq. (3). In the language of the algebraic Bethe ansatz[11,3] the scalar product is taken in the ''quantum'' (vertical) space, so one gets a 2×2 matrix in the ''auxiliary'' (horizontal) space:

$\langle\beta|R(z)|\alpha\rangle =$ 

The M operator (4) generates shifts of a time variable m. The ZC condition

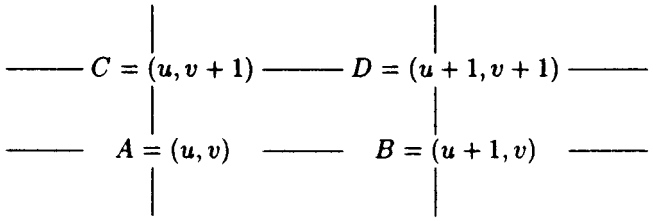
$$\mathcal{M}_{l+1,m}(z)\mathcal{L}_{l,m}(z)=\mathcal{L}_{l,m+1}(z)\mathcal{M}_{l,m}(z) \tag{6}$$

gives rise to the discrete soliton equations from Refs. 6 and 8.

The change of dynamical variables to the pair of vectors $|\alpha_l\rangle$, $|\beta_l\rangle$ plays a key role. Using the equations of motion of the discrete model, we show that (suitably normalized) components of the vectors $|\alpha_l\rangle$, $|\beta_l\rangle$ are τ *functions* (concerning τ functions see, e.g., Ref. 12).

In this paper we elaborate the simplest example — the lattice sine-Gordon (SG) model. There are two lattice versions of the classical SG model: the model on a space lattice with continuous time[5,13] and Hirota's SG equation on a space–time lattice.[8] They have common L operator. The M operators are given by Eq. (3) with the trigonometric classical r matrix for the former and by Eq. (4) for the latter, with $R(z)$ being the simplest trigonometric solution of the quantum Yang–Baxter equation (the R matrix of the XXZ spin chain).

2. By the SG model on a space–time lattice we mean the Faddeev–Volkov version[14,15] of Hirota's discrete SG equation.[8] This is a nonlinear equation for a function $\psi(u,v)$ on the 2D square lattice. Let



be an elementary cell of the u,v lattice. In this notation the equation reads

$$\nu\psi_C\psi_D-\nu\psi_A\psi_B=\mu(\psi_B\psi_D-\mu\psi_A\psi_C), \tag{7}$$

where μ,ν are constants. It contains both the Korteweg–de Vries (KdV) and sine-Gordon (SG) equations as different continuum limits. Equation (7) can be represented[14] as the ZC condition $L_{D\leftarrow B}(z;\nu)L_{B\leftarrow A}(z;\mu)=L_{D\leftarrow C}(z;\mu)L_{C\leftarrow A}(z;\nu)$ with the L matrix[14,16]

$$L_{B\leftarrow A}(z;\mu)=\begin{pmatrix} \mu\psi_B^{\frac12}\psi_A^{-\frac12} & z\psi_B^{-\frac12}\psi_A^{-\frac12} \\ z\psi_B^{\frac12}\psi_A^{\frac12} & \mu\psi_B^{-\frac12}\psi_A^{\frac12} \end{pmatrix}. \tag{8}$$

We call $l=\frac12(u+v)$, $m=\frac12(u-v)$ discrete space and time coordinates, respectively. Consider ''composite'' L and M operators generating shifts $A\to D$ and $C\to B$, respectively: $\hat{\mathcal{L}}_{D\leftarrow A}(z)=z^{-1}L_{D\leftarrow C}(z;\mu)L_{C\leftarrow A}(z;\nu)$, $\hat{\mathcal{M}}_{B\leftarrow C}(z)=z^{-1}(z^2-\nu^2)\times L_{B\leftarrow A}(z;\mu)[L_{C\leftarrow A}(z;\nu)]^{-1}$. From (8) we find:

$$\hat{\mathcal{L}}_{D\leftarrow A}(\mu z)=\begin{pmatrix} \mu z\psi_A^{\frac12}\psi_D^{-\frac12}+\nu z^{-1}\psi_D^{\frac12}\psi_A^{-\frac12} & \psi_C^{-1}\left(\mu\psi_D^{\frac12}\psi_A^{-\frac12}+\nu\psi_A^{\frac12}\psi_D^{-\frac12}\right) \\ \psi_C\left(\mu\psi_A^{\frac12}\psi_D^{-\frac12}+\nu\psi_D^{\frac12}\psi_A^{-\frac12}\right) & \mu z\psi_D^{\frac12}\psi_A^{-\frac12}+\nu z^{-1}\psi_A^{\frac12}\psi_D^{-\frac12} \end{pmatrix}, \tag{9}$$

$$\hat{\mathcal{M}}_{B\leftarrow C}(\mu z)=\begin{pmatrix} \mu z\psi_C^{\frac12}\psi_B^{-\frac12}-\nu z^{-1}\psi_B^{\frac12}\psi_C^{-\frac12} & \psi_A^{-1}\left(\mu\psi_B^{\frac12}\psi_C^{-\frac12}-\nu\psi_C^{\frac12}\psi_B^{-\frac12}\right) \\ \psi_A\left(\mu\psi_C^{\frac12}\psi_B^{-\frac12}-\nu\psi_B^{\frac12}\psi_C^{-\frac12}\right) & \mu z\psi_B^{\frac12}\psi_C^{-\frac12}-\nu z^{-1}\psi_C^{\frac12}\psi_B^{-\frac12} \end{pmatrix}. \tag{10}$$

The L operator of the lattice SG model with continuous time[5] at lth site is[a)]

$$\hat{\mathcal{L}}_l^{(IK)}(z)=\begin{pmatrix} z\chi_l+z^{-1}\chi_l^{-1} & s^{-\frac12}\varphi_l\pi_l \\ s^{-\frac12}\varphi_l\pi_l^{-1} & z\chi_l^{-1}+z^{-1}\chi_l \end{pmatrix}. \tag{11}$$

Here π_l and χ_l are exponentiated canonical variables, $\varphi_l=[1+s(\chi_l^2+\chi_l^{-2})]^{1/2}$, and s is a parameter. To identify the L operators (11) and (9), consider the composite fields

$\pi(u,v)=\psi^{1/2}(u+1,v)\psi^{1/2}(u,v+1)$, $\chi(u,v)=\psi^{1/2}(u,v)\psi^{-1/2}(u+1,v+1)$ and set $\pi_l=\pi(l,l)$, $\chi_l=\chi(l,l)$ at the constant time slice $m=0$. Identifying $s=\mu\nu(\mu^2+\nu^2)^{-1}$ and using Eq. (7), we conclude that $\hat{\mathcal{L}}_l^{(IK)}(z)=(\mu\nu)^{-1/2}\hat{\mathcal{L}}_l((\mu\nu)^{1/2}z)$. Here $\hat{\mathcal{L}}_l(z)\equiv\hat{\mathcal{L}}_{D_l\leftarrow A_l}(z)$, where $A_l=(l,l), D_l=(l+1,l+1)$. Similarly, we write $\hat{\mathcal{M}}_{\overline{B}_l\leftarrow A_l}(z)\equiv\hat{\mathcal{M}}_l(z)$, where $\overline{B_l}=(l+1,l-1)$. Then the discrete ZC condition acquires the form (6). The L operator $\hat{\mathcal{L}}_l^{(IK)}(z)$ has two degeneracy points $z_0^{\pm}=(\mu/\nu)^{\pm\frac{1}{2}}$ at which it is a projector (2) with the right-hand side expressed in terms of the field $\psi(u,v)$.

3. The idea of Hirota's approach[7] is to treat Eq. (7) as a consequence of 3-term bilinear equations for τ functions (see also Refs. 10 and 17). In the case at hand we need two τ functions: τ and $\hat{\tau}$. Set

$$\psi(u,v)=\frac{\hat{\tau}(u,v)}{\tau(u,v)} \tag{12}$$

then Eq. (7) follows from

$$(\nu-\mu)\hat{\tau}_A\tau_D=\nu\tau_B\hat{\tau}_C-\mu\hat{\tau}_B\tau_C, \quad (\nu-\mu)\tau_A\hat{\tau}_D=\nu\hat{\tau}_B\tau_C-\mu\tau_B\hat{\tau}_C. \tag{13}$$

The equivalent form of these equations,

$$(\nu+\mu)\tau_B\hat{\tau}_C=\mu\tau_A\hat{\tau}_D+\nu\hat{\tau}_A\tau_D, \quad (\nu+\mu)\hat{\tau}_B\tau_C=\mu\hat{\tau}_A\tau_D+\nu\tau_A\hat{\tau}_D, \tag{14}$$

is equally useful. At last, we point out the relation

$$\tau(u-1,v)\hat{\tau}(u+1,v)+\hat{\tau}(u-1,v)\tau(u+1,v)=2\tau(u,v)\hat{\tau}(u,v). \tag{15}$$

A few remarks are in order. Equations (13) form a part of the 2-reduced 2D Toda lattice hierarchy,[18] where μ, ν are *Miwa's variables*.[9] They play the role of inverse lattice spacings for the elementary discrete flows u,v. The lattice spacing in the m direction is then $(\mu\nu)^{-1}(\mu-\nu)$. Note that the u and v coordinate axes are in general not orthogonal to each other. In particular, as is seen from Eqs. (13), for $\mu=\nu$ one must *identify* u with v, and so the 2D lattice collapses to a 1D one. In this sense Eq. (15) follows from Eq. (14) for $\nu=\mu$.

4. We are ready to represent the M operator as a convolution of quantum R matrix with some vectors in its ''quantum'' space. Consider the quantum R matrices

$$R^{(\pm)}(z;q)=(a(z)\pm b(z))I\otimes I+(a(z)\mp b(z))\sigma_3\otimes\sigma_3+c(\sigma_1\otimes\sigma_1+\sigma_2\otimes\sigma_2), \tag{16}$$

where $a(z)=qz-q^{-1}z^{-1}$, $b(z)=z-z^{-1}$, $c=q-q^{-1}$, q is a ''quantum'' parameter, and z is the spectral parameter. The R matrices $R^{(+)}$ and $R^{(-)}$ differ by Drinfeld's twist. Both of them satisfy the quantum Yang–Baxter equation (in Sec. 1 we had $R(z)=R^{(+)}\times(z;q)$).

Let $|\alpha\rangle$, $|\beta\rangle$ be two vectors (see Eq. (2)) from the first (''quantum'') space. Consider the convolution $\langle\beta|R^{(\pm)}(z;q)|\alpha\rangle$ in the first space. This is a 2×2 matrix in the second (''auxiliary'') space:

$$\langle\beta|R^{(\pm)}(z;q)|\alpha\rangle=\begin{pmatrix}\beta^{(1)}\alpha^{(1)}a(z)\pm\beta^{(2)}\alpha^{(2)}b(z) & \beta^{(2)}\alpha^{(1)}c(z)\\ \beta^{(1)}\alpha^{(2)}c(z) & \pm\beta^{(1)}\alpha^{(1)}b(z)+\beta^{(2)}\alpha^{(2)}a(z)\end{pmatrix}. \tag{17}$$

Let us compare this with right-hand sides of Eqs. (9) and (10). To do that, we write elements of the L and M operators in terms of the τ functions (12) and after that use Eqs. (13) and (14) when necessary. The best result is achieved after the simple gauge transformation

$$\mathcal{L}_{A\leftarrow D}(z)=\left(\frac{\tau_D\hat{\tau}_D}{\tau_A\hat{\tau}_A}\right)^{1/2}\hat{\mathcal{L}}_{D\leftarrow A}(z),\quad \mathcal{M}_{B\leftarrow C}(z)=\left(\frac{\tau_B\hat{\tau}_B}{\tau_C\hat{\tau}_C}\right)^{1/2}\hat{\mathcal{M}}_{B\leftarrow C}(z). \tag{18}$$

Let us skip the details and present the final result. We set $\langle\alpha|=(\tau,\hat{\tau})$, $\langle\beta|=(\hat{\tau},\tau)$, and $q=\mu/\nu$. At the slice $m=0$ we have

$$\mathcal{L}_l(\mu z)=\frac{2\mu\nu}{\mu-\nu}\,\frac{\langle\beta_l|R^{(-)}(z;q)|\alpha_l\rangle}{\langle\beta_l|\alpha_{l-1}\rangle},\quad \mathcal{M}_l(\mu z)=\frac{2\mu\nu}{\mu+\nu}\,\frac{\langle\beta_l|R^{(+)}(z;q)|\check{\beta}_{l-1}\rangle}{\langle\beta_l|\alpha_{l-1}\rangle} \tag{19}$$

where the notation from the end of Sec. 2 is used. Up to the constant prefactors these formulas coincide with the ones announced in Sec. 1. The location of the vectors

$$|\alpha_l\rangle=\begin{pmatrix}\tau(l,l+1)\\ \hat{\tau}(l,l+1)\end{pmatrix}\quad\text{and}\quad |\beta_l\rangle=\begin{pmatrix}\hat{\tau}(l+1,l)\\ \tau(l+1,l)\end{pmatrix} \tag{20}$$

is shown in the first diagram, displayed below Eq. (5).

The normalization factor in Eq. (2) is equal to $\lambda_l=\mu\nu(\mu-\nu)^{-1}\tau(l,l)\hat{\tau}(l,l)$.

5. At last we show that the r-matrix formula (3) is a degenerate case of Eq. (4). A naive continuous-time limit would be $\nu\to\mu$, i.e., $q\to 1$, so that, in agreement with Eq. (3), we do get the r matrix. However, this would imply $\lim_{q\to 1}|\check{\beta}_l\rangle=|\alpha_l\rangle$ that is certainly wrong in general. The naive limit does not work, since the L operator itself varies as $\nu\to\mu$. In the correct limit the time lattice spacing must approach zero independently of μ, ν.

Let us introduce v' — another ''copy'' of the discrete flow v with Miwa's variable ν', so that we now have a 3D lattice. Equations of the type (13) are valid in the 2D sections $v'=$ const, $u=$ const, $v=$ const. Now we can let $\nu'\to\mu$ while leaving ν unchanged. We set $q'=\mu/\nu'=1+\varepsilon+O(\varepsilon^2)$, $\varepsilon\to 0$, where ε is the lattice spacing in the direction $m'=1/2(u-v')$. The discrete M operators are defined up to multiplication by a scalar function of z independent of the dynamical variables. It is convenient to normalize the M operators by $\mathcal{M}_l(z)=I$ at $\varepsilon=0$. Then the next term (of order ε) yields the continuous-time M operator. To find it, we expand in ε the discrete M operator $\mathcal{M}_{B'_l\leftarrow C'_l}(z)$, which generates the shift $(l-1,l,1)\to(l,l,0)$ on the 3D lattice with coordinates (u,v,v').

The second diagram, displayed below Eq. (6), shows the u,v' section. The coordinates of the vertices are: $A_l'=(l-1,l,0)$, $B_l'=A_l=(l,l,0)$, $C_l'=(l-1,l,1)$, $D_l'=(l,l,1)$. The point C_l' tends to the point $B_l'=A_l$ as $\nu'\to\mu$, so the parallelogram collapses to the u axis. We have: $\mathcal{M}_{B_l'\leftarrow C_l'}(z)=I+\varepsilon M_l(z)+O(\varepsilon^2)$, where

$$M_l(\mu z)=\frac{1}{z-z^{-1}}\begin{pmatrix}\frac{1}{2}(z+z^{-1})\dfrac{\tau(l-1,l)\hat{\tau}(l+1,l)}{\tau(l,l)\hat{\tau}(l,l)} & \dfrac{\tau(l-1,l)\tau(l+1,l)}{\tau(l,l)\hat{\tau}(l,l)}\\ \dfrac{\hat{\tau}(l-1,l)\hat{\tau}(l+1,l)}{\tau(l,l)\hat{\tau}(l,l)} & \frac{1}{2}(z+z^{-1})\dfrac{\hat{\tau}(l-1,l)\tau(l+1,l)}{\tau(l,l)\hat{\tau}(l,l)}\end{pmatrix}. \tag{21}$$

The r matrix is $r(z)=\lim_{\varepsilon\to 0}\varepsilon^{-1}[(z+z^{-1})^{-1}R^{(+)}(z;q')-I\otimes I]$, and so

$$r(z)=\frac{1}{2(z-z^{-1})}[(z+z^{-1})I\otimes I+2\sigma_1\otimes\sigma_1+2\sigma_2\otimes\sigma_2+(z+z^{-1})\sigma_3\otimes\sigma_3]. \tag{22}$$

Comparing with (21), we get Eq. (3) with the r matrix (22).

6. The main result of this work is the R matrix representation (19) of the local L–M pair for the classical SG model in discrete space–time. In our opinion, the very fact that the typical quantum R matrix naturally arises in a purely classical problem is important and interesting by itself. It would be desirable to clarify a connection with the quantum Yang–Baxter equation (which already arose in purely classical problems in a different context.[19,20]) We should stress that the ''quantum'' parameter q of the R matrix in our context is related to the mass parameter and the lattice spacing of the classical model.

I thank S. Kharchev and P. Wiegmann for steadfast interest in this work, very helpful discussions, and critical remarks. Discussions with O. Lipan, I. Krichever and A. Volkov are also gratefully acknowledged. This work was supported in part by RFBR Grant No. 97-02-19085.

[a)]We take the L operator from Ref. 5 and multiply it by σ_2 from the left so as to deal with Eq. (7) rather than Hirota's equation.

[1] L. Faddeev and L. Takhtadjan, *Hamiltonian Methods in the Theory of Solitons*, Springer, 1987.
[2] V. Zakharov, S. Manakov, S. Novikov and L. Pitaevskiy, *Theory of Solitons. The Inverse Problem Method,* Moscow, Nauka, 1980.
[3] E. K. Sklyanin, Zap. Nauchn. Semin. LOMI **95** 55 (1980).
[4] V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin, *Quantum Inverse Scattering Method and Correlation Functions*, Cambridge University Press, 1993.
[5] A. G. Izergin and V. E. Korepin, Lett. Math. Phys. **5**, 199 (1981); Nucl. Phys. B **205**, 401 (1982).
[6] R. Hirota, J. Phys. Soc. Jpn. **43**, 1424 (1977).
[7] R. Hirota, J. Phys. Soc. Jpn. **43**, 2074 (1977); J. Phys. Soc. Jpn. **50**, 3785 (1981).
[8] R. Hirota, J. Phys. Soc. Jpn. **43**, 2070 (1977).
[9] T. Miwa, Proc. Jap. Acad. **58**, Ser. A, 9 (1982).
[10] E. Date, M. Jimbo, and T. Miwa, J. Phys. Soc. Jpn. **51**, 4116 (1982).
[11] L. Faddeev and L. Takhtadjan, Usp. Mat. Nauk **34**(5), 14 (1979).
[12] M. Jimbo and T. Miwa, Publ. RIMS, Kyoto Univ. **19**, 943 (1983); G. Segal and G. Wilson, Publ. IHES **61**, 5 (1985).

[13] V. Tarasov, Zap. Nauchn. Semin. LOMI **120**, 173 (1982).
[14] L. Faddeev and A. Volkov, Lett. Math. Phys. **32**, 125 (1994).
[15] A. Volkov, Univ. of Uppsala preprint (1995), hep-th/9509024.
[16] L. D. Faddeev, *Lectures at E. Fermi Summer School*, Varenna 1994, hep-th/9406196.
[17] A. Zabrodin, preprint ITEP-TH-10/97, solv-int/9704001.
[18] K. Ueno and K. Takasaki, Adv. Studies Pure Math. **4**, 1 (1984).
[19] E. K. Sklyanin, Zap. Nauchn. Semin. LOMI **146**, 119 (1985).
[20] A. Weinstein and P. Xu, Commun. Math. Phys. **148**, 309 (1992).

Published in English in the original Russian journal. Edited by Steve Torstveit.