How to determine an effective potential for a variable cosmological term

A. A. Starobinsky

Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia

It is shown that if a variable cosmological term in the present Universe is described by a scalar field with minimal coupling to gravity and with some phenomenological self-interaction potential $V(\varphi)$, then this potential can be unambiguously determined from the following observational data: either from the behavior of density perturbations in dustlike matter component as a function of redshift (given the Hubble constant additionally), or from the luminosity distance as a function of redshift (given the present density of dustlike matter in terms of the critical value). © *1998 American Institute of Physics.*
[S0021-3640(98)00122-4]

PACS numbers: 98.80.Cq, 95.35.+d

It has been known for many years that the flat Friedmann–Robertson–Walker (FRW) cosmological model with cold dark matter (CDM), a positive cosmological constant $\Lambda > 0$ ($\Omega_0 + \Omega_\Lambda = 1$) and an approximately flat (or a Harrison–Zeldovich-like, $n_S \approx 1$) spectrum of primordial scalar (adiabatic) perturbations fits the observational data better and has a larger admissible region of the parameters (H_0, Ω_0) than any other cosmological model, both with inflationary and noninflationary initial conditions (see, e.g., Refs. 1 and 2). Here H_0 is the Hubble constant, $\Omega_0 = 8\pi G\rho_m/3H_0^2$ includes baryons and (mainly) nonbaryonic dark matter, $\Omega_\Lambda \equiv \Lambda/3H_0^2$, and the light velocity $c = 1$. This conclusion was based on the following arguments: a) the relation between H_0 and the age of the Universe t_0, b) the fact that the observed mass/luminosity ratio never leads to values greater than $\Omega_0 \sim 0.4$ up to supercluster scales, c) comparison of the cosmic microwave background (CMB) temperature anisotropies, the power spectra of density and velocity matter perturbations, and the present abundance of galaxy clusters with predictions of cosmological models with inflationary initial conditions; d) the observed values of ρ_b/ρ_m in rich galaxy clusters taken together with the range for the present baryon density ρ_b admitted by the theory of primordial (Big Bang) nucleosynthesis. I don't include gravitational lensing tests (e.g., the number of lensed quasars) here, since conclusions based on them are less definite at present; however, the most recent reconsideration[3] has also led to a low value $\Omega_0 \sim 0.3$.

During last year two new pieces of strong evidence for $\Omega_0 < 1$ have appeared. The first (historically) of them is based on the evolution of the abundance of rich galaxy

clusters with redshift z (Ref. 4; see also the more recent paper by Eke *et al.*,[5] where the value $\Omega_0 \approx 0.5 \pm 0.2$ (1σ uncertainty) is given). Still, it should be noted that some doubts have already been cast on the validity of the conclusion that $\Omega_0 = 1$ is really excluded.[6] Much better observational data expected in the near future will help to resolve this dilemma unambiguously. The second, completely independent argument for $\Omega_0 = 0.2-0.4$ follows from direct observations of supernovae type Ia (SNIa) explosions at high redshifts up to $z \sim 1$.[7] On the other hand, no direct evidence for a negative spatial curvature of the Universe (i.e., for the open FRW model) has been found. Just the opposite — the latest CMB constraints (based mainly on the results of the Saskatoon and CAT experiments),[8] the galaxy abundance at high redshifts,[9] and the most recent analysis of the SNIa data in terms of an effective equation of state of a component adding Ω_0 to unity[10] strongly disfavor the open CDM model without a positive cosmological constant. Of course, the possibility of having *both* a positive cosmological constant and spatial curvature of any sign is not yet excluded, but, according to the ''Occam's razor'' principle, it would be desirable not to introduce one more basic novel feature of the Universe (spatial curvature) without conclusive observational evidence. In any case, in spite of many theoretical and experimental attempts to exorcize it, a Λ term is back again.

It is clear that the introduction of a cosmological constant requires new and completely unknown physics in the region of ultralow energies. Solutions with a cosmological constant occur in such fundamental theories as supergravity and M theory. However, this cosmological constant is always negative and very large. As compared to such a basic ''vacuum'' state, a very small and positive cosmological constant allowed in the present Universe may be thought as corresponding to the energy density ε_Λ of a highly excited (though still very symmetric) ''background'' state,.so it need not be very ''fundamental.'' But then it is natural to omit the assumption that it should be exactly constant. In this case the name ''a cosmological term'' (or a Λ term) is more relevant for it, so I shall use this one below. The principal difference between two kinds of nonbaryonic dark matter — dustlike CDM and a Λ term — is that the latter one is not gravitationally clustered up to scales $\sim 30 \ h^{-1}$ or more (otherwise we would return to the problem why Ω_0 observed from gravitational clustering is not equal to unity). Here $h = H_0/100$ km$\cdot$s^{-1}Mpc^{-1}.

On the other hand, there exists a well-known strong argument showing that a Λ term cannot change with time as fast as the matter density ρ_m and the Ricci tensor (i.e., $\propto t^{-2}$) during the matter-dominated stage (for redshifts $z < 4 \times 10^4 h^2$). Really, if $\varepsilon_\Lambda \propto \rho_m$, so that $\Omega_\Lambda = $ const, then matter density perturbations in the CDM+baryon component grow as

$$\delta \equiv \left(\frac{\delta\rho}{\rho} \right)_m \propto t^\alpha \propto (1+z)^{-3\alpha/2}, \quad \alpha = \frac{\sqrt{25 - 24\Omega_\Lambda} - 1}{6}.$$

As a consequence, the total growth of perturbations Δ since the time of equality of matter and radiation energy densities up to the present moment is less than in the absence of the Λ term. If $\Omega_\Lambda \ll 1$, then $\Delta(\Omega_\Lambda) = \Delta(0)(1 - (6.4 + 2 \ln h)\Omega_\Lambda)$. Since the parameters of viable cosmological models are so tightly constrained that Δ may not be reduced by more than about twice, this type of a Λ term cannot account for more than ~ 0.1 of the critical energy density (see Ref. 11 for detailed investigation confirming this conclusion). This, unfortunately, precludes a natural explanation of the present Λ term with $\Omega_\Lambda = (0.5-0.8)$

in terms of ''compensation'' mechanisms[12] or exponential potentials with sufficiently large exponents;[13] in other words, a Λ term cannot be produced by an exactly ''tracker'' field, as was recently proposed in Ref. 14.

A natural and simple description of a variable Λ term is just that which was so successively used to construct the simplest versions of the inflationary scenario, namely, a scalar field with some interaction potential $V(\varphi)$ minimally coupled to the Einstein gravity. Such an approach, though phenomenological, is nevertheless more consistent and fundamental than a commonly used attempt to describe a Λ term by a barotropic ideal fluid with some equation of state. The latter approach cannot be made internally consistent in the case of negative pressure, which is implied by observations;[10] in particular, it generally leads to imaginary values of the sound velocity. On the contrary, no such problems arise using the scalar field description (this scalar field is called the Λ field below). Of course, its effective mass $|m_\varphi^2| = |d^2V/d\varphi^2|$ should be very small to avoid gravitational clustering of this field in galaxies, clusters, and superclusters. To make a Λ term slowly varying, we assume that $|m_\varphi| \sim H_0 \sim 10^{-33}$ eV or less (though this condition may be relaxed). Models with a time-dependent Λ term were introduced more than ten years ago,[15] and different potentials $V(\varphi)$ (all inspired by inflationary models) have been considered: exponential,[13,16,11,17] inverse power-law,[18] power-law,[19] and cosine.[20,17]

However, it is clear that since we know essentially nothing about physics at such energies, there exists no preferred theoretical candidate for $V(\varphi)$. In this case, it is more natural to go from observations to theory, and to determine an effective phenomenological potential $V(\varphi)$ from observational data. The two new tests mentioned above are the most suitable for this purpose. Indeed, using the cluster abundance $n(z)$ determined from observations and assuming Gaussian statistics of the initial perturbations (the latter follows from the paradigm of one-field inflation, and it is in agreement with other observational data), it is possible to determine a *linear* density perturbation in the CDM+baryon dustlike component $\delta(z)$ for a fixed comoving scale $R \sim 8(1+z)^{-1}h^{-1}$ Mpc up to $z \sim 1$, either by using the Press–Schechter approximation or by direct numerical simulations of nonlinear gravitational instability in the expanding Universe. $\delta(z)$ can be also determined from observation of gravitational clustering (in particular, of the galaxy–galaxy correlation function) as a function of z. On the other hand, observations of SNe at different z yield the luminosity distance $D_L(z)$ through the standard astronomical expression $m = M + 5 \log D_L + 25$, where m is the observed magnitude, M is the absolute magnitude, and D_L is measured in Mpc.

The aim of the present letter is to show how to determine $V(\varphi)$ from either $\delta(z)$ or $D_L(z)$, and to investigate what additional information is necessary for an unambiguous solution of this problem in both cases. The idea has already been announced by the author in Refs. 21 and 22; now the details are given.

The derivation of $V(\varphi)$ consists of two steps. First, the Hubble parameter $H \equiv \dot{a}/a = H(z)$ is determined. Here $a(t)$ is the FRW scale factor, $1+z \equiv a_0/a$, the dot means d/dt and the index 0 denotes the present value of a corresponding quantity (in particular, $H(t_0) = H(z=0) = H_0$). In the case of SNe, the first step is almost trivial since the textbook expression for D_L reads:

$$D_L(z) = a_0(\eta_0 - \eta)(1+z), \quad \eta = \int_0^t \frac{dt}{a(t)}. \tag{1}$$

Therefore,

$$H(z) = \frac{da}{a^2 d\eta} = -(a_0 \eta')^{-1} = \left[\left(\frac{D_L(z)}{1+z} \right)' \right]^{-1}. \tag{2}$$

Here and below, a prime denotes the derivative with respect to z. Thus, $D_L(z)$ defines $H(z)$ uniquely.

More calculations are required to find $H(z)$ from $\delta(z)$. The system of background equations for the system under consideration is:

$$H^2 = \frac{8\pi G}{3} \left(\rho_m + \frac{\dot{\varphi}^2}{2} + V \right), \quad \rho_m = \frac{3\Omega_0 H_0^2 a_0^3}{8\pi G a^3}, \tag{3}$$

$$\ddot{\varphi} + 3H\dot{\varphi} + \frac{dV}{d\varphi} = 0, \tag{4}$$

$$\dot{H} = -4\pi G(\rho_m + \dot{\varphi}^2). \tag{5}$$

Equation (5) is actually a consequence of the other two equations.

We consider a perturbed FRW background which metric, in the longitudinal gauge (LG), has the form:

$$ds^2 = (1 + 2\Phi)dt^2 - a^2(t)(1 + 2\Psi)\delta_{lm}dx^l dx^m; \quad l,m = 1,2,3. \tag{6}$$

The system of equations for scalar perturbations reads (the spatial dependence $\exp(ik_l x^l)$, $k_l k^l \equiv k^2$ is assumed):

$$\Phi = \Psi = \dot{v}, \quad \dot{\delta} = -\frac{k^2}{a^2}v + 3(\ddot{v} + H\dot{v} + \dot{H}v), \tag{7}$$

$$\dot{\Phi} + H\Phi = 4\pi G(\rho_m v + \dot{\varphi}\delta\varphi), \tag{8}$$

$$\left(-\frac{k^2}{a^2} + 4\pi G\dot{\varphi}^2 \right)\Phi = 4\pi G\left(\rho_m \delta + \dot{\varphi}\dot{\delta\varphi} + 3H\dot{\varphi}\delta\varphi + \frac{dV}{d\varphi}\delta\varphi \right), \tag{9}$$

$$\ddot{\delta\varphi} + 3H\dot{\delta\varphi} + \left(\frac{k^2}{a^2} + \frac{d^2V}{d\varphi^2} \right)\delta\varphi = 4\dot{\varphi}\dot{\Phi} - 2\frac{dV}{d\varphi}\Phi. \tag{10}$$

Equation (10) is a consequence of the other equations. Here v and $\delta\varphi$ are, correspondingly, a velocity potential of a dustlike matter peculiar velocity and a Λ-field perturbation in LG, and δ is a *comoving* fractional matter density perturbation (in this case, it coincides with $(\delta\rho/\rho)_m$ in the synchronous gauge). In fact, all these perturbed quantities are gauge-invariant.

Now let us take a comoving wavelength $\lambda = k/a(t)$ which is much smaller than the Hubble radius $H^{-1}(t)$ up to redshifts $z \sim 5$. This corresponds to $\lambda \ll 2000\ h^{-1}$ Mpc at present. Then, from Eq. (10),

$$\delta\varphi \approx \frac{a^2}{k^2}\left(4\dot{\varphi}\dot{\Phi}-2\frac{dV}{d\varphi}\Phi\right), \quad |\dot{\varphi}\delta\varphi|\sim\left|\frac{dV}{d\varphi}\delta\varphi\right|\sim\frac{a^2H^4}{Gk^2}|\Phi|\ll\rho_m|\delta|. \tag{11}$$

Therefore, the Λ field is practically unclustered at the scale involved. Now the last of Eqs. (7) and Eq. (9) may be simplified to:

$$\dot{\delta}=-\frac{k^2}{a^2}v, \quad -\frac{k^2}{a^2}\Phi=4\pi G\rho_m\delta. \tag{12}$$

Combining this with the first of Eqs. (7), we return to a well-known equation for δ in the absence of the Λ field:

$$\ddot{\delta}+2H\dot{\delta}-4\pi G\rho_m\delta=0. \tag{13}$$

It is not possible to solve this equation analytically for an arbitrary $V(\varphi)$. Remarkably, the inverse dynamical problem, i.e., the determination of $H(a)$ given $\delta(a)$, is solvable. After changing the argument in Eq. (13) from t to a $[d/dt=aH(d/da)]$, we get a first order linear differential equation for $H^2(a)$:

$$a^2\frac{d\delta}{da}\frac{dH^2}{da}+2\left(a^2\frac{d^2\delta}{da^2}+3a\frac{d\delta}{da}\right)H^2=\frac{3\Omega_0H_0^2a_0^3\delta}{a^3}. \tag{14}$$

The solution is:

$$H^2=\frac{3\Omega_0H_0^2a_0^3}{a^6}\left(\frac{d\delta}{da}\right)^{-2}\int_0^a a\delta\frac{d\delta}{da}da=3\Omega_0H_0^2\frac{(1+z)^2}{\delta'^2}\int_z^\infty\frac{\delta|\delta'|}{1+z}dz. \tag{15}$$

Putting $z=0$ in this expression for H, we arrive to the expression of Ω_0 through $\delta(z)$:

$$\Omega_0=\delta'^2(0)\left(3\int_0^\infty\frac{\delta|\delta'|}{1+z}dz\right)^{-1}. \tag{16}$$

Of course, observations of gravitational clustering can hardly provide the function $\delta(z)$ for too large z (say, for $z>5$). However, $\delta(z)$ in the integrands in Eqs. (15) and (16) may be well approximated by its $\Omega_0=1$ behavior (i.e., $\delta\propto(1+z)^{-1}$) already for $z>(2-3)$. If massive neutrinos are present, one should use here the expression with α written above and with Ω_Λ replaced by Ω_ν/Ω_0 (it is assumed that ρ_m includes massive neutrinos, too).

Finally, using Eq. (16), Eq. (15) can be represented in a more convenient form:

$$\frac{H^2(z)}{H^2(0)}=\frac{(1+z)^2\delta'^2(0)}{\delta'^2(z)}-3\Omega_0\frac{(1+z)^2}{\delta'^2(z)}\int_0^z\frac{\delta|\delta'|}{1+z}dz. \tag{17}$$

Thus, $\delta(z)$ uniquely defines the ratio $H(z)/H_0$. Of course, appearance of derivatives of $\delta(z)$ in these formulas shows that sufficiently clean data are necessary, but one may expect that such data will soon appear. Let us recall also that, for $\Lambda\equiv$const ($V(\varphi)\equiv$const), we have

$$H^2(z)=H_0^2(1-\Omega_0+\Omega_0(1+z)^3), \quad q_0\equiv-1+\left(\frac{d\ln H}{d\ln(1+z)}\right)_{z=0}=\frac{3}{2}\Omega_0-1, \tag{18}$$

where q_0 is the acceleration parameter.

The second step — the derivation of $V(\varphi)$ from $H(a)$ — is very simple. One has to rewrite Eqs. (3) and (5) in terms of a and take their linear combinations:

$$8\pi G V(\varphi) = aH\frac{dH}{da} + 3H^2 - \frac{3}{2}\Omega_0 H_0^2\left(\frac{a_0}{a}\right)^3,$$

$$4\pi G a^2 H^2\left(\frac{d\varphi}{da}\right)^2 = -aH\frac{dH}{da} - \frac{3}{2}\Omega_0 H_0^2\left(\frac{a_0}{a}\right)^3, \tag{19}$$

and then exclude a from these equations.

Therefore, the model of a Λ term considered in this paper can account for *any* observed forms of $D_L(z)$ and $\delta(z)$ which, in turn, can be transformed into a corresponding effective potential $V(\varphi)$ of the Λ field. The only condition is that the functions $H(z)$ obtained in these two independent ways should coincide within observational errors. $D_L(z)$ uniquely determines $V(\varphi)$ if Ω_0 is given additionally [the latter is required at the second step, in Eqs. (19)]. $\delta(z)$ uniquely determines $V(\varphi)$ up to the factor H_0^2, which has to be given additionally to fix an overall amplitude. Observational tests which can disprove this model do exist. In particular, a contribution to large-angle $\Delta T/T$ CMB temperature anisotropy due to the integrated (or, nonlocal) Sachs–Wolfe effect presents a possibility for distinguishing the model from more complicated models, e.g., with non-minimal coupling of the Λ field to gravity or to CDM. However, the latter test is not an easy one, since this contribution is rather small and is partially masked by cosmic variance.

[1] L. A. Kofman and A. A. Starobinsky, Pis'ma Astron. Zh. **11**, 643 (1985) [Sov. Astron. Lett. **11**, 271 (1985)]; L. A. Kofman, N. Yu. Gnedin, and N. A. Bahcall, Astrophys. J. **413**, 1 (1993); J. P. Ostriker and P. J. Steinhardt, Nature (London) **377**, 600 (1995); J. S. Bagla, T. Padmanabhan, and J. V. Narlikar, Comments. Astrophys. **18**, 275 (1996) (http://xxx.lanl.gov/abs/astro-ph/9511102).

[2] A. A. Starobinsky, in *Cosmoparticle Physics. 1. Proceedings of the 1st International Conference on Cosmoparticle Physics ''Cosmion-94''*, Moscow, 5–14 Dec. 1994, edited by M. Yu. Khlopov, M. E. Prokhorov, A. A. Starobinsky, and J. Tran Thanh Van, Frontiers, 1996, p. 141 (http://xxx.lanl.gov/abs/astro-ph/9603074).

[3] M. Chiba and Y. Yoshii, http://xxx.lanl.gov/abs/astro-ph/9808321.

[4] N. A. Bahcall, X. Fan, and R. Cen, Astrophys. J. Lett. **485**, L53 (1997); X. Fan, N. A. Bahcall, and R. Cen, Astrophys. J. Lett. **490**, L123 (1997).

[5] V. R. Eke, S. Cole, C. S. Frenk, and J. P. Henry, MNRAS, in press (1998) (http://xxx.lanl.gov/abs/astro-ph/9802350).

[6] R. Sadat, A. Blanchard, and J. Oukbir, Astron. Astrophys. **329**, 21 (1997); P. Viana and A. R. Liddle, MNRAS, in press (1998) (http://xxx.lanl.gov/abs/astro-ph/9803244); A. Blanchard, J. G. Bartlett, and R. Sadat, C. R. Acad. Sci., in press (1998) (http://xxx.lanl.gov/abs/astro-ph/9809182).

[7] P. M. Garnavich, R. P. Kirshner, P. Challis *et al.*, Astrophys. J. Lett. **493**, L53 (1998); S. Perlmutter, G. Aldering, M. Della Valle *et al.*, Nature (London) **391**, 51 (1998); A. G. Riess, A. V. Filippenko, P. Challis *et al.*, Astrophys. J. in press (1998) (http://xxx.lanl.gov/abs/astro-ph/9805201).

[8] C. H. Lineweaver, Astrophys. J. Lett., submitted (1998) (http://xxx.lanl.gov/abs/astro-ph/9805326).

[9] J. A. Peacock, http://xxx.lanl.gov/abs/astro-ph/9805208.

[10] P. M. Garnavich, S. Jha, P. Challis *et al.*, Astrophys. J., in press (1998) (http://xxx.lanl.gov/abs/astro-ph/9806396).

[11] P. G. Ferreira and M. Joyce, Phys. Rev. Lett. **79**, 4740 (1997); Phys. Rev. D **57**, 4686 (1998); E. J. Copeland, A. R. Liddle, and D. Wands, Phys. Rev. D **58**, 023503 (1998).

[12] A. D. Dolgov, in *The Very Early Universe*, edited by G. W. Gibbons, S. W. Hawking, and S. T. C. Siklos, Cambridge University Press (1983), p. 449; A. D. Dolgov, JETP Lett. **41**, 345 (1985).

[13] C. Wetterich, Nucl. Phys. B **302**, 668 (1988); Astron. Astrophys. **301**, 321 (1995).

[14] I. Zlatev, L. Wang, and P. J. Steinhardt, http://xxx.lanl.gov/abs/astro-ph/9807002.

[15] M. Özer and M. O. Taha, Phys. Lett. B **171**, 363 (1986); Nucl. Phys. B **287**, 776 (1987); M. Reuter and C. Wetterich, Phys. Lett. B **188**, 38 (1987); K. Freese, F. C. Adams, J. A. Frieman, and E. Mottola, Nucl. Phys. B **287**, 797 (1987); M. Gasperini, Phys. Lett. B **194**, 347 (1987); N. Weiss, Phys. Lett. B **197**, 42 (1987).

[16] B. Ratra and P. J. E. Peebles, Phys. Rev. D **37**, 3406 (1988).

[17] P. T. P. Viana and A. R. Liddle, Phys. Rev. D **57**, 674 (1998).

[18] P. J. E. Peebles and B. Ratra, Astrophys. J. Lett. **325**, L17 (1988).

[19] N. Weiss, Phys. Rev. D **39**, 1517 (1989).

[20] J. A. Frieman, C. T. Hill, A. Stebbins, and I. Waga, Phys. Rev. Lett. **75**, 2077 (1995); K. Coble, S. Dodelson, and J. Frieman, Phys. Rev. D **55**, 1851 (1997).

[21] A. A. Starobinsky, *Large Scale Structure: Tracks and Traces, Proceedings of the 12th Potsdam Cosmology Workshop*, Potsdam, Sept. 15–20, 1997, edited by V. Müller *et al.*, Singapore, World Scientific (1998) (http://xxx.lanl.gov/abs/astro-ph/9808152).

[22] A. A. Starobinsky, *15th International Conference on General Relativity and Gravitation*, Dec. 16–21, 1997, Abstracts of plenary lectures and contributed papers, IUCAA (1997), p. 180.

Resonance behavior of a $K_S K_S$ system near the mass 1775 MeV

B. P. Barkov, O. N. Baloshin, N. V. Batalova, V. V. Vladimirskiĭ,
V. K. Grigor'ev, O. N. Erofeeva, Yu. V. Katinov, I. Ya. Korol'kov, V. I. Lisin,
V. N. Luzin, V. N. Nozdrachev, V. V. Sokolovskiĭ,[a] G. D. Tikhomirov,
S. A. Uzunyan, and Yu. P. Shkurenko

Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia

The preliminary results of an investigation of a system of two K_S mesons in the mass interval 1600–1950 MeV are reported. The events were obtained on a 6-m magnetic spark spectrometer at ITEP in $\pi^- p$ interactions at 40 GeV, using a neutral trigger which suppressed both charged particles and γ rays. A peak of width $\simeq 30$ MeV with statistical significance not lower than six standard deviations is observed with momentum transfer selection $|t| > 0.23$ GeV^2 near the mass 1775 MeV of the $K_S K_S$ system. The observed phenomena can be interpreted as the existence of one resonance with the indicated parameters, or two narrower resonances. In the latter case, their masses are 1768 ± 1.5 and 1787 ± 1.5 MeV. The widths of these states are comparable to the mass resolution of the spectrometer (~ 5 MeV). Estimates of the product $\sigma \cdot \mathrm{BR}(K_S K_S)$ give ~ 1.5 and 2.5 nb, respectively, for the first and second states. © *1998 American Institute of Physics.*
[S0021-3640(98)00222-9]

PACS numbers: 14.40.Aq, 13.60.Le

A peak centered at 1775 MeV has been observed in the 1600–1950 MeV region of invariant masses of the $K_S K_S$ system using the 6-m spectrometer at ITEP. The width ~ 30 MeV of this feature is almost an order of magnitude smaller than the typical width of resonances in this mass range. We investigated the $K_S K_S$ system with momentum transfers greater than a certain value, which was fixed during the analysis process. The purpose of such selection was to suppress the dominant contribution of one-pion exchange. In previous work, a narrow resonance of width $\Gamma < 70$ MeV was observed in a system of two η mesons in the indicated mass interval using a GAMS spectrometer with momentum transfer selection similar to ours.[1–4]

In the present letter we report the preliminary results of an analysis of a narrow peak observed in the $K_S K_S$ system. It was found that the peak can be described by one state or two narrower states, the widths of both resonances in the latter case being of the order of the instrumental mass resolution of the spectrometer.

The statistics of K_S-meson pair production, which is investigated in this work on the basis of $\sim 30\,000$ events, was obtained by detecting the reaction

$$\pi^- p \to K_S K_S n \tag{1}$$

using the 6-m magnetic spark spectrometer (MSS) at ITEP at momentum 40 GeV. The spectrometer was placed in the beam of the accelerator at the Institute of High-Energy Physics (Protvino). A detailed description of the apparatus and the results of the investigation of the $K_S K_S$ system in the MSS at ITEP is contained in Refs. 5, 6 and 7–9, respectively. A trigger apparatus was used to detect the reaction (1). The main element of the trigger are veto counters, surrounding a liquid-hydrogen target. The counters form a double shielding layer around the target. To suppress not only charged particles but also γ rays emanating from the target, lead converters were placed between the counters. The detected events include events for which (their fraction $<20\%$) a baryon and one or several pions are produced instead of a neutron at the lower vertex:

$$\pi^- p \to K_S K_S (n\pi^0, \dots, p\pi^-, \dots). \tag{2}$$

These events are due to the nonideal efficiency of the veto counters and the counting losses of the electronics.

K_S mesons are identified from their decay into the pair $\pi^+\pi^-$. The detection efficiency of the $K_S K_S$ system for the indicated decay mode in the mass range 1600–1950 is $\sim 40\%$. The directions of the momenta of the π mesons are measured with an accuracy of ~ 0.3 mrad. The mass width of the K_S meson used to calibrate the apparatus is 10 MeV (full width at half maximum). Fitting of the tracks to the intersection at the vertices of the forks and to the tabulated value of the $K_S K_S$ meson mass greatly improves the accuracy of the calculation of the physical parameters. As a result, the error in determining the effective mass of two K_S mesons in the region up to 1800 MeV is no more than 5 MeV.

The kinematic variables used here to analyze a system of two K_S mesons are the effective mass M_{KK} of a pair of K_S mesons, the squared missing mass MM^2 (the squared mass of the particles which are not detected by the spectrometer), and the modulus of the squared 4-momentum transfer $|t|$.

Figure 1 shows the effective-mass distribution of two K_S mesons with t selection ($0.23 < |t| < 0.60$ GeV^2) and with the limitation on the squared missing mass $MM^2 < 8$ GeV^2. A peak whose statistical significance is higher than six standard deviations is observed near 1775 MeV: At the maximum there are 110 events, the expected number being 53 events. The statistical significance here and below was determined from an estimate of the statistical error according to mass intervals next to the peak. In the present case one standard deviation was taken to be $\sqrt{53}$. In what follows, we shall designate the peak by $X(1775)$. We note that in the mass spectrum there are no statistically significant deviations in the channel 1775 ± 15 MeV for small momentum transfer ($|t| < 0.23$ GeV^2).

As we have said, along with the reaction (1), events belonging to the reaction (2) are detected. The contribution of these processes can be judged from the distributions of the number of events over the squared missing mass MM^2 to the produced pair $K_S K_S$. Figure 2 shows the distribution over MM^2 for the region $M_{KK} = 1600$–1950 MeV (the curve shows events with $|t| < 0.23$ GeV^2 selection, while the dots with error bars show events with $0.23 < |t| < 0.60$ GeV^2). In the distribution represented by the curve, the events are concentrated in a region with a width of the order of 2 GeV^2 around the neutron mass. The width is determined mainly by the variance of the beam momentum.
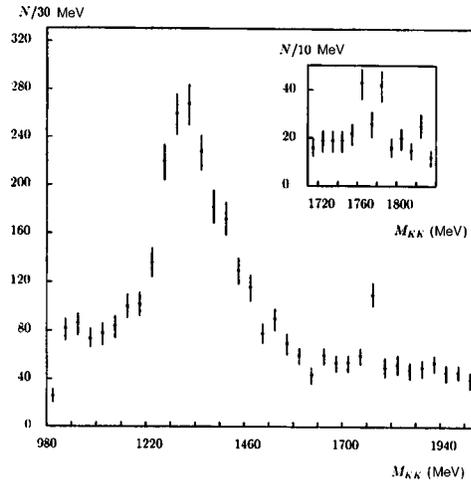
FIG. 1. Distribution of events over the invariant mass of the $K_S K_S$ system with a 30 MeV step with selections $0.23 < |t| < 0.60$ GeV^2, $MM^2 < 8$ GeV^2. Inset: Distribution with a 10 MeV step.

For large transfers (dots with error bars) the distribution is shifted in the direction of the mass of the isobar $\Delta(1232)$. Moreover, an appreciable fraction of events corresponding to heavier missing masses ($MM^2 > 3.0$ GeV^2) is present. For small transfers the number of events which do not fit into the neutron peak does not exceed 5% of the total number, and for large transfers it is about 30%. This fraction can be decreased by introducing strict MM^2 selection, but the feature observed in the present work also appears in events belonging to the reaction (2). The soft limit which we use on the missing mass ($MM^2 < 8$ GeV^2) cuts off no more than 5% of the events.
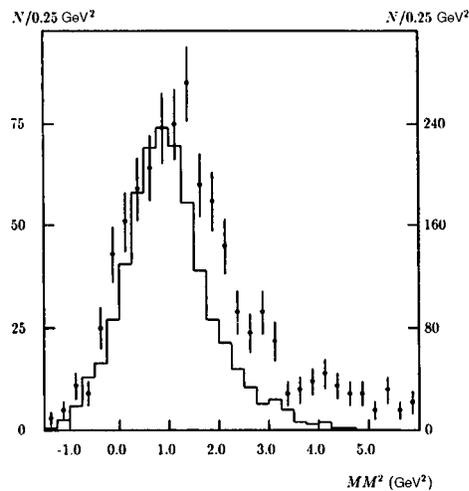
FIG. 2. Distributions of the number of events over the squared missing mass to a system of two kaons whose effective mass lies in the interval 1600–1950 MeV. The histogram shows events with $|t| < 0.23$ GeV^2 (right-hand scale); the dots show events with $|t| > 0.23$ GeV^2 (left-hand scale).
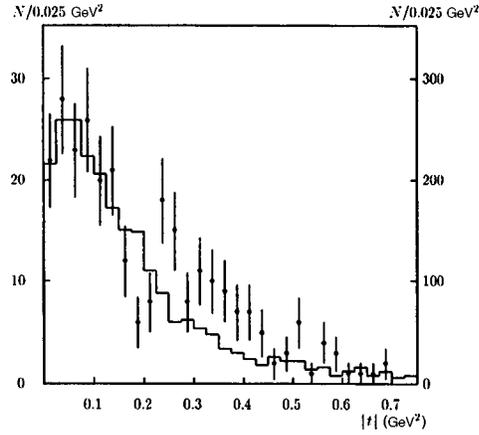
FIG. 3. Distributions of the number of events over the squared momentum transfer $|t|$. Histogram — events whose mass M_{KK} lies in the bands 1600–1760 and 1790–1950 MeV (right-hand scale). Dots — events in the band 1760–1790 MeV (left-hand scale).

Figure 3 shows the $|t|$ dependence of the yield of $K_S K_S$ pairs for events whose effective mass lies in the band 1775 ± 15 MeV (dots with error bars) as well as for events whose mass lies in the neighboring bands 1600–1760 and 1790–1950 MeV (histogram). Comparing the two distributions, it is evident that for $0 < |t| < 0.15$ GeV^2 their shapes match well. For larger momentum transfers the distribution for the band 1760–1790 MeV is irregular. For $|t|$ in the range from 0.23 to 0.45 GeV^2, as compared with the histogram (adjacent bands), this distribution contains an excess of events which within the statistical error is numerically close to the excess of events above background at the peak 1775 MeV (see Fig. 1). The lower limit of the excess with respect to $|t|$ is determined quite sharply. The limiting value $|t| = 0.23$ GeV^2 for $|t|$-selected events is optimal from the standpoint of the best signal/background ratio. The upper limit of the momentum transfer is unimportant, since fewer than 15% of the events fall into the region of large transfers.

Figure 4 shows the mass distribution of a system of two K_S mesons with a 5 MeV step. One can see from this figure that the experimental distribution in the mass interval 1760–1790 MeV has a quite complicated structure, which can be described by one resonance or two resonances. To obtain a quantitative answer to the question of the number of peaks, the experimental data were fit using the maximum likelihood method (MLM) and the least-squares method (LSM). The background was described by a constant, and one or two Gaussian functions were used to describe the resonance behavior.

The functional used in the MLM has the form

$$\int_{M_{KK}} \epsilon(M_{KK}) F(P, M_{KK}) dM_{KK} - \ln L, \tag{3}$$

where

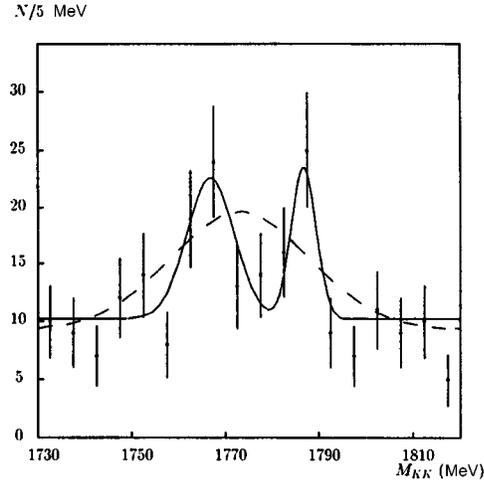$$L = \prod_{i=1}^{N} F(P, M_{KK}^i)$$

$N/5$ MeV

FIG. 4. Results of a maximum likelihood fit of the mass distribution: Dots — experimental data, dashed curve — result of a fit with one resonance, solid line — two resonances.

is the maximum likelihood function (the product extends over all N events) and $\epsilon(M_{KK})$ is the detection efficiency for two K_S mesons. The parameters P include the squares of the background amplitude A, the amplitude B_k of the production of resonances, and the parameters M_k and δ_k of the Gaussian function. The probability density F is given by the expression

$$F(P;M_{KK})=A+\sum_{k=1}^{2} B_k G(M_k,\delta_k;M_{KK}),\tag{4}$$

where G is a Gaussian function.

The parameter χ^2 was calculated according to the formula

$$\chi^2=-2\ln L+\text{const}\tag{5}$$

using the value of L obtained by minimizing expression (3) (see, for example, Ref. 10). The constant was chosen so that the MLM and LSM values of χ^2 would be the same for the data described without resonances. The values of χ^2 obtained by both methods differ from one another by not more than three units. However, the MLM gives a more accurate estimate of the parameters, and it is free of the conventions which are inherent in the LSM — the choice of the channel width and the zero point of the reading does not affect the result of the fit. At the same time, the LSM makes it possible to estimate not only the relative but also the absolute reliability of a particular hypothesis. The fact that the values of χ^2 obtained by both methods are close to one another makes it possible to compare the MLM value of χ^2 with the number of degrees of freedom corresponding to the LSM.

The results obtained by fitting the experimental data by the MLM without including resonance states, with one resonance, and with two resonances are presented in Table I. The first row of the table gives the value of χ^2 for the case without resonances. For the fit with one resonance the maximum of the solution coincides with the center of the feature (1775 MeV) and embraces the entire phenomenon ($\delta\simeq 15$ MeV). The results of

TABLE I.

Type of fit	$\chi^2/N_{\text{deg. free.}}$	M_k, MeV	δ_k, MeV	B_k/A
No resonances	44/19	–	–	0
With one resonance	28/16	1774.5 ± 2.0	14.3 ± 2.0	1.15 ± 0.2
With two resonances	15/13	1767.5 ± 1.7	5.0 ± 2.5	1.34 ± 0.3
		1786.7 ± 0.7	2.5 ± 0.5	2.36 ± 0.8

the fit of the data with two resonances are presented in the last two rows of the table. The values of χ^2 obtained in solutions with resonances do not permit choosing between them.

Figure 4 illustrates the solution with one resonance (dashed curve) and two resonances (solid curve).

Let us summarize the results obtained. A feature of width ~ 30 MeV was observed near the mass 1775 MeV in a system of two K_S mesons produced in the reactions (1) and (2) at $|t| > 0.23$ GeV^2. The statistical significance of the existence of the feature is not lower than six standard deviations. The observed phenomenon can be interpreted as the existence of one resonance or two narrower resonances. In the first case the parameters of the resonance are $M = 1775 \pm 1.5$ MeV, $\Gamma = 30 \pm 7$ MeV, and $\sigma \cdot \text{BR}(K_S K_S) = 5.0 \pm 1.0 \pm 2.0$ nb. For the case of two resonances the masses are 1768 ± 1.5 and 1787 ± 1.5 MeV for the lighter and heavier states. For both of them the estimated widths $\Gamma < 5.0$ MeV are comparable to the mass resolution of the spectrometer, and the products $\sigma \cdot \text{BR}(K_S K_S)$ are $1.5 \pm 0.3 \pm 1.0$ nb and $2.5 \pm 0.5 \pm 1.5$ nb, respectively.

[a] e-mail: sokolovsky@vitep5.itep.ru

[1] D. Alde, F. Binon, K. Bricman *et al.*, JETP Lett. **44**, 567 (1986).
[2] D. Alde, F. Binon, C. Bricman *et al.*, Phys. Lett. B **182**, 105 (1986).
[3] D. Alde, F. Binon, K. Bricman *et al.*, Yad. Fiz. **54**, 745 (1991) [Sov. J. Nucl. Phys. **54**, 451, (1991)].
[4] D. Alde, F. Binon, C. Bricman *et al.*, Phys. Lett. B **284**, 757 (1992).
[5] B. V. Bolonkin, O. N. Baloshin, A. M. Blagorodov *et al.*, Preprint No. 86 [in Russian], ITEP, 1973.
[6] B. V. Bolonkin, V. V. Vladimirskiĭ, A. P. Grishin *et al.*, Preprint No. 154 [in Russian], ITEP, 1981.
[7] B. V. Bolonkin, S. K. Bloshenko, V. V. Vladimirskiĭ *et al.*, Yad. Fiz. **46**, 799 (1987) [Sov. J. Nucl. Phys. **46**, 451 (1987)].
[8] B. V. Bolonkin, S. K. Bloshenko, V. V. Vladimirsky *et al.*, Nucl. Phys. B **309**, 426 (1988).
[9] V. K. Grigor'ev, O. N. Baloshin, B. P. Barkov *et al.*, Yad. Fiz. **59**, 2187 (1996) [Phys. At. Nucl. **59**, 2105 (1996)].
[10] Particle Data Group, Phys. Rev. D **54**, 26 (1996).

Translated by M. E. Alferieff

Observation of two-step excitation of photoluminescence in silicon nanostructures

L. A. Golovan',[a)] A. A. Goncharov, V. Yu. Timoshenko, A. P. Shkurinov, P. K. Kashkarov, and N. I. Koroteev

M. V. Lomonosov Moscow State University, Department of Physics, and International Laser Center, 119899 Moscow, Russia

Efficient visible-range photoluminescence with photon energy higher than the photon energy of the exciting radiation is observed in nanostructures of porous silicon subjected to heat treatment in vacuum. The photoluminescence intensity is found to be virtually identical for cw and femtosecond excitation by Ti:sapphire laser radiation with the same average power. The results can be explained by a two-step cascade photoluminescence excitation process in which optical passivation of defects of the dangling silicon bond type occurs. © *1998 American Institute of Physics.* [S0021-3640(98)00322-3]

PACS numbers: 78.66.Vs, 78.55.Ap, 61.46.+w

Electrochemical treatment of a single-crystal silicon wafer in a solution based on hydrofluoric acid produces an up to 100 μm thick porous layer on the surface of the sample. Under certain conditions of formation, the layer of porous silicon (PS) contains silicon filaments and clusters with 1–5 nm cross sections. As a result of the quantum size effect, the band gap E_g in PS is larger than the value characteristic for bulk silicon $E_{g0} = 1.1$ eV (at $T = 300$ K). Efficient visible-range room-temperature photoluminescence (PL) is observed in such a material (see, for example, the review in Ref. 1).

Ordinarily, PL is excited in PS by radiation with photon energy $h\nu > E_g$. Some electron–hole pairs produced in the process are bound in excitons, whose radiative annihilation explains the visible-range PL in unoxidized porous silicon.[2] This process competes with nonradiative (or radiative in the IR range) recombination on surface defects, for example, dangling silicon bond type centers.[1,2] In as-prepared PS samples these bonds are partially or completely passivated by hydrogen atoms, which results in a low recombination rate on surface defects.[1] Vacuum heating of PS at temperatures above 650 K removes the hydrogen from the surface of nanoclusters. This gives rise to a large number of dangling silicon bonds. Nonradiative recombination of charge carriers on these defects becomes dominant, and for this reason in such PS samples there is virtually no visible-range PL.[1,2]

In the present work it was established experimentally that visible-range PL in PS nanostructures with a high density of surface defects by illumination with photons with energy $h\nu < E_g$. The observed effect can be explained by a cascade excitation process in

which defects, whose levels lie in the band gap of PS, participate.

The PS samples were obtained by electrochemical etching of single crystal silicon wafers in a 1:2 solution of HF in ethanol. The substrate consisted of *p*-type silicon, obtained by the method of zone melting, with (111) orientation and resistivity $\sim 1\Omega \cdot$ cm. Two types of PS samples were prepared. One type was formed in 6 min with a current density 30 mA/cm^2 and the other in 3 min with current density 120 mA/cm^2. The porosities of the PS layers were $\sim 70\%$ and 85%. The porous layers were about 4 μm thick. Immediately after preparation, the PS samples were placed in a vacuum cell with working pressure $\sim 10^{-4}$ Pa. Both as-prepared samples and samples subjected to heat treatment in vacuum at 700 K for 45 min were used.

The PL was excited by radiation from a N$_2$ laser ($\lambda = 337$ nm, $\tau = 10$ ns, energy density per pulse 0.5 mJ/cm^2, pulse repetition frequency 100 Hz) and a Ti:sapphire laser. The latter laser operated in the cw mode and a femtosecond pulse mode ($\tau = 100$ fs) with a repetition frequency of 100 MHz. The laser could be wavelength-tuned in the range $\lambda_{exc} = 755-790$ nm.[3] The Ti:sapphire laser radiation was focused into a spot with a diameter of the order of 10 μm. The intensity of the radiation in both modes could reach ~ 10 kW/cm^2. The PL was detected with monochromators and photomultipliers (MDR-23 and FÉU-136 with excitation by the N$_2$ laser and MSD-1 and a Hamatsu R106UH for excitation by the Ti:sapphire laser).

We shall now present the experimental results. When freshly prepared samples were illuminated by N$_2$ laser radiation, PL was observed in the range 600–750 nm for a sample with 70% porosity and 500–650 nm for a sample with 85% porosity. In both types of samples the vacuum heat treatment result in complete quenching of the PL excited by the N$_2$-laser radiation.

Visible-range PL was observed in PS samples with 70% porosity, both freshly prepared and after heat treatment in vacuum, excited by Ti:sapphire laser radiation. In the first case, however, the PL was several orders of magnitude weaker. After the PS was heated its PL was visible with the naked eye. For PS with 85% porosity the weak PL observed under the same conditions for a freshly prepared sample disappeared after vacuum heat treatment. The PL spectra of PS with 70% porosity excited by Ti:sapphire laser radiation are presented in Figs. 1 and 2. They consist of wide bands in the wavelength interval 500–600 nm. As one can see from Fig. 1, the intensity and spectrum of the PL were close for excitation by cw and femtosecond lasing regimes with the same average intensity ($I_{exc} = 7$ kW/cm^2). At the same time, a strong nonlinear dependence of the PL signal on the average excitation intensity is observed (Fig. 2). An appreciable PL signal was detected with I_{exc} above the threshold value 4 kW/cm^2 (inset in Fig. 2).

It was found that the intensity of the PL of PS depends strongly on the wavelength of the exciting radiation (Fig. 3). At the same time, the spectral composition of the PL changed very little. The PL signal with maximum intensity was recorded for a laser radiation wavelength of 770 nm.

Since intense light fluxes were used to excite PL with the Ti:sapphire laser radiation, we shall first estimate the possible thermal effects. The absorption coefficient for light with $\lambda_{exc} = 755-790$ nm in the experimental PS layers does not exceed 10^2 cm^{-1}.[1] Since thin (4 μm) layers were used, most (up to 96%) of the radiation was absorbed in the silicon substrate. As shown in Ref. 4, the heating of the surface of a silicon single crystal
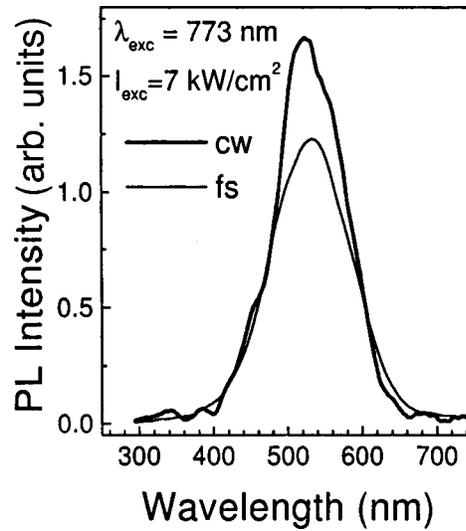
FIG. 1. The PL spectra of PS samples excited by cw and femtosecond pulse Ti:sapphire laser radiation with the same average intensity $I_{\mathrm{exc}} = 7$ kW/cm^2.

by ~ 10 kW/cm^2 Ti:sapphire laser radiation did not exceed several degrees. Therefore the heating of the PS layers on the silicon substrate side can be neglected. The temperature of a porous layer which has absorbed laser radiation can be estimated using the thermal conductivity of high-porosity PS $(1.55 \times 10^{-3}$ W/cm).[5] According to such estimates, the temperature increase does not exceed 150 K. Therefore the corresponding thermal radia-
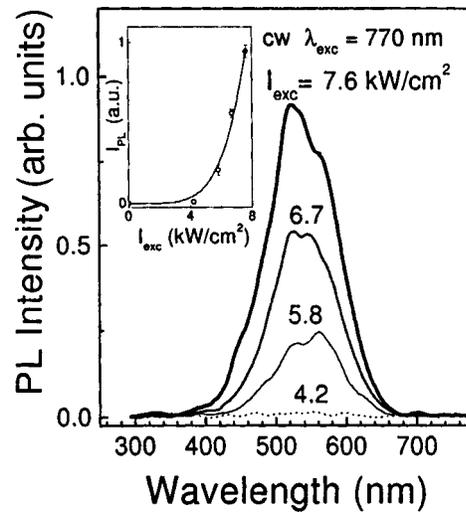

FIG. 2. PL spectra of PS excited by cw Ti:sapphire laser radiation at wavelength 770 nm with intensities ranging from 4.2 to 7.6 kW/cm^2 (numbers above the spectra). Inset: PL intensity I_{PL} at the band maximum versus the excitation intensity I_{exc}.
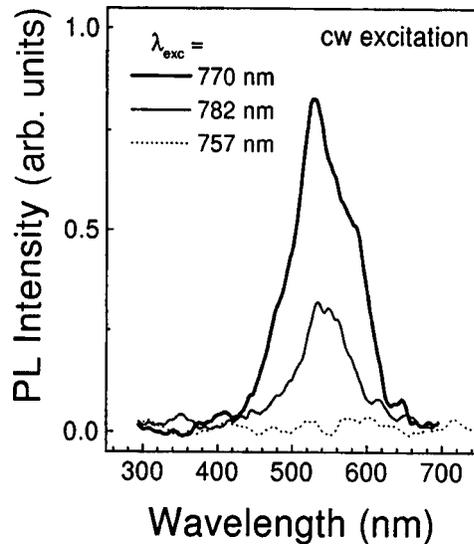
FIG. 3. PL spectra of PS excited by cw Ti:sapphire laser radiation with various wavelengths: 770 nm (I_{exc} = 8.2 kW/cm^2), 782 nm (I_{exc} = 8.2 kW/cm^2), and 757 nm (I_{exc} = 6.4 kW/cm^2).

tion cannot make a large contribution to the detected PL signal. An additional argument against the observed radiation in PS being of a thermal nature is that the position of the maximum of the PL band does not depend on the intensity of the exciting radiation (see Fig. 2). Moreover, the resonance character of the dependence of the PL intensity on the radiation wavelength does not agree with the assumption of thermal radiation (see Fig. 3).

Photoluminescence similar to that obtained by excitation of PS with N_2-laser radiation has been studied well. It can be explained by radiative annihilation of excitons which are excited by interband absorption in filament or cluster type silicon nanostructures, with transverse dimensions of the order of 1–2 nm.[1,6] Quenching of PL as a result of vacuum heat treatment likewise corresponds well to the published data.[1]

Obviously, defects due to dangling silicon bonds on the surface of nanoclusters play a large role in the excitation of PL in heat-treated PS samples by Ti:sapphire laser radiation. This shows that the PL efficiency is appreciably higher for vacuum-heated samples. We note that under the same heat-treatment conditions the density of dangling bonds in PS samples reached 10^{18} cm^{-3} (Ref. 7). Apparently, the presence of defect levels in the band gap of PS increases the probability of absorption of two laser photons, which results in the appearance of a pair of free nonequilibrium charge carriers. This process is not two-photon absorption, since the intensity of the PL which arises is virtually identical for femtosecond-pulse and cw excitation (see Fig. 1). Therefore the PL intensity is determined by the average intensity of the laser radiation and not by its peak value in the case of pumping with femtosecond pulses. A two-step cascade process with a change in the populations of the defect levels in the band gap of PS probably occurs. At the first stage of the cascade process an electron is excited optically into a defect level and is trapped by this center. At the second stage, excitation of a trapped electron into the conduction band occurs as a result of absorption of the next optical photon.

The process of cascade excitation of PL depends on the pump wavelength (see Fig. 3). This could be due to a resonance mechanism of absorption of optical photons. The above-noted absence of PL in high-porosity samples agrees with this supposition. Indeed, as the nanoclusters become smaller, the band gap and the distance between the band edge and the defect level increase.[6] Therefore the conditions for resonance absorption are no longer satisfied. Since the optimal wavelength for excitation of PL is known (Fig. 3), the energy of a valence band–defect level transition can estimated. We obtain 1.6 eV. The maximum in the PL spectrum corresponds to $h\nu \simeq 2.3$ eV. Assuming the PL to be of excitonic character, and taking into account the binding energy of an electron and a hole,[6] we obtain $E_6 \simeq 2.5$ eV.

An additional factor determining the high efficiency of PL in heat-treated PS samples excited by Ti:sapphire laser radiation is that the rate of nonradiative recombination on dangling silicon bond type defects is lower than for excitation by N_2-laser radiation ($h\nu > E_g$). Probably, the recombination centers on the surface of silicon nanostructures on which nonradiative recombination with interband absorption occurs are entrained into the cascade process of excitation of PL by radiation with $h\nu < E_g$. The substantial filling of the levels of these defects with electrons in the process of resonance absorption decreases the probability of capture of an electron from the conduction band, and these centers are thereby excluded from the recombination process. A unique kind of *optical passivation* of recombination centers occurs. The sharp dependence of the PL intensity on the pump intensity (inset in Fig. 2) is explained by the simultaneous action of cascade absorption and optical passivation mechanisms.

In summary, it was established experimentally that efficient excitation of visible-range luminescence occurs when vacuum heat-treated PS is illuminated by Ti:sapphire laser radiation. The excitation of photoluminescence is most likely due to a two-step resonance cascade process with participation of dangling silicon bond type defects. To clarify the details of this process it is necessary to perform additional theoretical and experimental investigations, including time-resolved measurements.

a)e-mail: leo@ofme.phys.msu.su

[1] A. G. Gullis, L. T. Canham, and P. D. J. Calcott, J. Appl. Phys. **82**, 909 (1997).
[2] P. K. Pashkarov, B. V. Kamenev, E. A. Konstantinova *et al.*, Usp. Fiz. Nauk **168**, 577 (1998).
[3] A. A. Angeluts, N. I. Koroteeva, S. A. Magnitskiĭ *et al.*, Prib. Tekh. Éksp. **8**, 94 (1998).
[4] J. I. Dadap, X. F. Hu, N. M. Russel *et al.*, IEEE J. Sel. Top. Quantum Electron. **1**, 1145 (1995).
[5] A. N. Obraztsov, H. Okushi, H. Watanabe, and V. Yu. Timoshenko, Phys. Status Solidi B **203**, 565 (1997).
[6] G. C. John and V. A. Singh, Phys. Rep. **263**, 93 (1995).
[7] E. A. Konstantinova, P. K. Kashkarov, and V. Yu. Timoshenko, Phys. Low-Dimens. Semicond. Struct. **12**, 127 (1995).

Translated by M. E. Alferieff

Charge transfer states in GeO$_2$-doped silicate fiber-optic waveguides and their role in second-harmonic generation

B. P. Antonyuk,[a)] V. N. Denisov, and B. N. Mavrin

Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

The concentration dependence of the broad-band luminescence arising in GeO$_2$-doped silicate fiber-optic waveguides is measured and interpreted. The spectra obtained show that electronic excitations of a new type that are absent in pure silicate glass — charge transfer excitons — arise in the doped system. Under the action of light an electron can be transferred both from a Ge center into the host and between Ge centers. Self-organization of these excitations in the field of a light wave (orientational ordering of their dipole moments) results in the appearance of a macroscopic electric field that destroys the initial centrosymmetry of the system and allows second-harmonic generation. © *1998 American Institute of Physics.* [S0021-3640(98)00422-8]

PACS numbers: 42.81.Qb, 42.65.Ky, 78.55.Hx

Weak second-harmonic generation (at the level of Raman scattering signals) exists in all fiber-optic waveguides.[1,2] Recent experiments[3] have shown that the frequency-doubling efficiency is proportional to the surface area of the waveguide. This means that weak generation occurs in a surface layer, where the centrosymmetry is broken and this process is allowed. Osterberg and Margulis have observed second-harmonic generation with ～10% efficiency in a GeO$_2$-doped silicate fiber-optic waveguide after transmission of Nd:YAG laser radiation for 5 h.[4] Such a high doubling efficiency attests to the fact that in this case the centrosymmetry was broken in the entire waveguide. This phenomenon has attracted the attention of many investigators.[5–9] In the first theory[5] the idea of ordering of defects was advanced. In all other theories the idea of asymmetric photoelectron emission from Ge centers was used. These theories use free electron and hole states excited by multiphoton processes, which make it possible to overcome the 8–10 eV ''gap'' existing in these systems. We performed experiments on Raman and hyper-Raman scattering of light in silicate fiber-optic waveguides doped with GeO$_2$ to molar concentrations of 5%, 10%, and 29%, and we compared their spectra with the corresponding spectra of pure silicate glass in order to determine the characteristic features that appear when GeO$_2$ is added (self-organization of the system, resulting in second-harmonic generation, has been observed in just such doped waveguides). The hyper-Raman scattering spectra were excited by radiation at the fundamental frequency of the
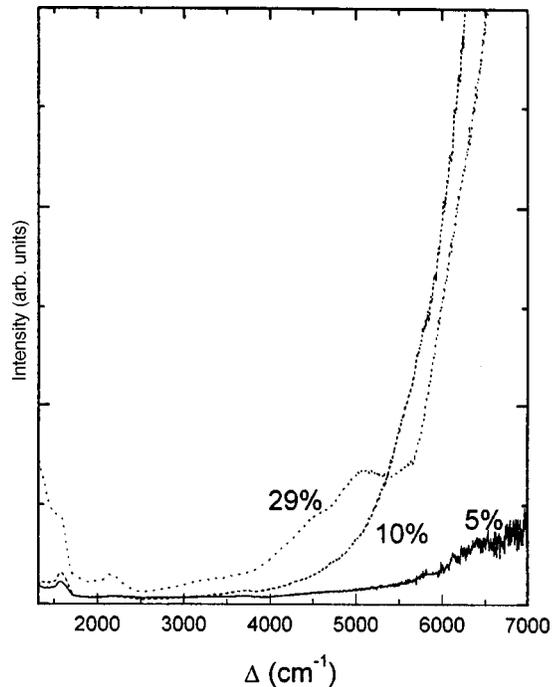
FIG. 1. Luminescence spectra of silicate fiber-optic waveguides doped with GeO_2 to molar concentrations of 5, 10, and 29%; Δ — red shift from the exciting line $\lambda = 488$ nm.

Nd:YAG laser ($\lambda = 1064$ nm); the Raman scattering spectra were observed with excitation by the second harmonic ($\lambda = 532$ nm) and by Ar^+-laser lines ($\lambda = 514.5$ nm and $\lambda = 488$ nm). We observed that GeO_2 impurities change only the vibrational spectra of hyper-Raman scattering. This means that the IR line $\lambda = 1064$ nm does not excite any new states in the doped waveguide. However, the lines $\lambda = 532$ nm, $\lambda = 514.5$ nm, and $\lambda = 488$ nm excite new wide luminescence bands in GeO_2-doped fiber-optic waveguides that are completely absent in the pure silicate glass. We checked that the new states are excited by a single photon of the monochromatic field of the indicated lines.

The luminescence spectra were recorded with a home-built spectrometer with triple monochromatization and a multichannel detector, consisting of an amplifying tube and a vidicon. An objective lens ($\times 40$) focused the laser beam into a fiber-optic waveguide, the entrance end of which could be displaced for adjustment in three mutually perpendicular directions. All waveguides were 1 m long and had a 12 μm in diameter core. The power of the radiation leaving a waveguide did not exceed 9 mW. The radiation was focused onto the slit of a spectrometer, which made it possible to detect the luminescence spectrum up to 750 nm. Since our detector records simultaneously the spectral region $2500-3000$ cm^{-1}, we detected the spectrum by parts, matching up their overlapping sections. The change in the spectral sensitivity of the spectrometer was corrected by comparing with a band-lamp spectrum with a known intensity distribution. The luminescence spectra were normalized to the intensity of the radiation leaving the waveguide.
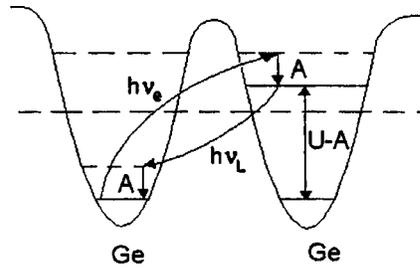
FIG. 2. Photoinduced transitions of an electron between Ge centers: the energy of exciting photon $h\nu_e$ ≈ 2.54 eV$\approx U$, the energy of luminescence photon $h\nu_L \approx 1.67$ eV, the Stokes shift $A \approx 1/2(h\nu_e - h\nu_L)$ ≈ 0.45 eV, and ϵ_F is the Fermi energy.

The spectra obtained in the spectral region above 523 nm are shown in Fig. 1. The luminescence spectrum below 523 nm overlaps with intense Raman scattering lines. Luminescence in this region was first observed in Ref. 10 and then in Ref. 11. We attribute to second-order processes the 528 nm (1570 cm^{-1}) band which is present in all spectra. As one can see from Fig. 1, additions of GeO$_2$ lead to the appearance of wideband (~ 1 eV) luminescence, excited by a single photon of green light ($\lambda = 488$ nm, $h\nu_e = 2.54$ eV). Since the ''gap'' in the silicate glass is 8–10 eV, this luminescence corresponds to recombination of local electron and hole states. The wide spectrum (~ 1 eV) attests to the fact that the particles make transitions between different potential wells with different energy levels. Under the monochromatic excitation employed electronic transitions between the levels of the same trap would give discrete levels that could be broadened by electron-phonon interaction up to $\sim h\nu_D \sim 10^{-2}$ eV (ν_D is the Debye frequency) but by no means not up to the observed width of 1 eV.

As one can see from Fig. 1, the luminescence intensity in the region $\nu < 5500$ cm^{-1} grows approximately linearly with the GeO$_2$ concentration. This attests to the fact that the centers participating in this luminescence are independent of one another. It is natural to attribute this luminescence to transitions of an electron from a Ge center into the host (absorption) and back (recombination with emission of a luminescence photon). Comparing the spectra for low (5%) and high (10%, 29%) concentrations reveals superlinear growth of the luminescence intensity in the region $\nu > 5500$ cm^{-1} as a function of the concentration. This shows that the centers giving this radiation interact with one another. It is natural to attribute the radiation to transitions between Ge centers. Under excitation an electron is transferred from one Ge center to another and then recombines after relaxation with a Stokes energy shift A, after which a second relaxation follows (Fig. 2). In the ground state there is one electron on each Ge center, while in the excited state a localized hole is present on one center and two electrons are present on another. The Coulomb interaction U of electrons on the same Ge center makes the dominant contribution to the excitation energy (Fig. 2). This gives the estimate $U \approx h\nu_e \approx 2.54$ eV. The experimental energy of a luminescence photon in this band is $h\nu_L \approx 1.67$ eV ($\lambda = 740$ nm, the shift of the band from the excitation line is 7000 cm^{-1}), so that $A = 1/2(h\nu_e - h\nu_L) \approx 0.45$ eV. Since $A \gg h\nu_D$, in the present case we have strong electron-phonon coupling, characteristic precisely for charge-transfer excitons to whose energy the Coulomb interaction of the separated electrons and holes, which depends

strongly on the electron–hole distance (the derivatives of the energy of the state with respect to the displacements of the atoms are large), makes the dominant contribution.

As one can see from Fig. 1, an increase of the GeO_2 concentration from 10% to 29% leads to the appearance of luminescence in the band under discussion $\nu > 5500$ cm^{-1}. Apparently, at such concentrations electron and hole migration along Ge centers, as a result of which relaxation of the excitation energy occurs, becomes important. In this case, less energy is released on recombination and the released energy can be carried off by phonons. This activates nonradiative recombination, which decreases the quantum yield of luminescence.

The excited charge transfer states, which are excited by a single photon of green light, play a special role in the theory of self-organization in fiber-optic waveguides, leading to second-harmonic generation.[11] Orientational ordering of their dipole moments leads to the appearance of macroscopic polarization and therefore a macroscopic electric field, which allows second-harmonic generation. The states observed in the present work serve as the basis of a new approach to the problem of self-organization. The model arising employs only localized electron and hole states and only single-phonon transitions between them.

Our preliminary investigations show that it is possible to explain on the basis of this approach all the rich experimental data accumulated over 12 years of investigations and never adequately explained on the basis of the models mentioned above.

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

[1] Y. Fujii, B. S. Kawasaki, K. O. Hill, and D. C. Johnson, Opt. Lett. **5**, 48 (1980).
[2] Y. Ohmori and Y. Sasaki, Appl. Phys. Lett. **39**, 466 (1981).
[3] V. B. Podobedov, J. Raman Spectrosc. **27**, 731 (1996).
[4] U. Osterberg and W. Margulis, Opt. Lett. **11**, 516 (1986).
[5] Q. H. Stolen and H. W. K. Tom, Opt. Lett. **12**, 585 (1987).
[6] E. M. Dianov, P. G. Kazanskii, and D. Y. Stepanov, Sov. J. Quantum Electron. **19**, 575 (1989).
[7] E. M. Dianov and D. S. Starodubov, Opt. Fiber Technol. **1**, 3 (1994).
[8] N. B. Baranova, A. N. Chudinov, and B. Y. Zel'dovich, Opt. Commun. **79**, 116 (1990).
[9] D. Z. Anderson, V. Mizrahi, and J. E. Sipe, Opt. Lett. **16**, 796 (1991).
[10] D. M. Krol, R. M. Atkins, and P. J. Lemaire in *Proceedings of the International Workshop on Photoinduced Self-Organization Effects in Optical Fibers*, 10–11 May 1991, Quebec City.
[11] B. P. Antonyuk, V. B. Antonyuk, S. F. Musichenko, and V. B. Podobedov, Phys. Lett. A **213**, 297 (1996).

Translated by M. E. Alferieff

Processes occurring in cluster plasmas and cluster beams

B. M. Smirnov

Institute for High Temperatures, Russian Academy of Sciences, 127412 Moscow, Russia

A method of generating a cluster beam of a refractory metal is analyzed. Clusters are formed in a high-pressure arc discharge plasma in which the refractory metal is an additive to the buffer gas. The conditions for cluster formation to occur in the intermediate region of the discharge are found, and the cluster instability, as a result of which all of the refractory metal collects in the form of clusters in this zone of the discharge, is studied. The processes studied are the basis for of a method of generating intense cluster beams. © *1998 American Institute of Physics.* [S0021-3640(98)00522-2]

PACS numbers: 39.10.+j, 36.40.−c, 52.80.Mg

A cluster plasma is a unique physical object, consisting of a dense buffer-gas plasma and trace impurity of a refractory metal whose atoms form clusters. It is convenient to use for the plasma medium the positive column of a high-pressure arc discharge.[1] The metal can be introduced into such a plasma in the form of a chemical compound. The compounds of the metal and its vapor can be spatially separated over the cross section of the discharge because of the high temperature gradients and the weakness of the transport processes (on account of the high density of the buffer gas), which prevents mixing of the different components of the metal. However, even though the transport processes are slow, the metal atoms can escape to the walls of the discharge tube. A process of chemical regeneration[2–5] of the refractory metal is used to prevent the attachment of atoms and clusters to the walls. Then, the buffer gas contains, along with the metal, a halogen as an additive, so that the gas component MX_k (M and X are atoms of the metal and halogen) forms near the walls, while the metal in the hot region of the discharge exists in the form of clusters or atomic vapor.[6] In what follows, we shall study the processes occurring in such a gas-discharge plasma.

Under the conditions considered, exchange between clusters and the gas component of the metal occurs as a result of diffusion of the components through a transitional zone separating the region containing clusters from the region containing the gas component. In parallel with this, clusters grow and vaporize efficiently in the region where they form. The characteristic cluster size can be determined by the competition between these processes. However, as clusters increase in size, cluster growth accelerates, while the diffusion process leading to the displacement of the clusters slows down. This results in instability, as a result of which the clusters remain stuck in the region where they grow, while the characteristic cluster size increases with time. As a result, the refractory metal

collects in the form of growing clusters in the cluster region of the arc discharge.

The cluster plasma under study can be used to generate an intense cluster beam. Ordinarily, cluster beams are obtained by allowing atomic vapor to expand through a nozzle.[7,8] The clusters grow in the beam as a result atoms attaching to them, while the beam is cooled as a result of expansion. The buffer gas in an expanding beam is used to convert the atomic vapor completely into clusters, while additional charging of the clusters makes it possible to control the beam. In a cluster plasma the process leading to the formation of charged clusters occurs automatically, so that charged clusters can be extracted from the plasma in the form of a beam. We shall assume that the plasma is allowed to escape from the cluster zone of the discharge through an opening, after which an electric field separates the charged clusters in the form of an autonomous beam from the plasma. A cluster beam thereby form as a result of a collection of processes occurring in the cluster plasma and the plasma flow. On account of the nonequilibrium nature of these processes, the possibility of formation of a cluster beam in the scheme considered depends on various parameters of the system. In this letter these processes are analyzed and the conditions for realizing the proposed scheme for generating a cluster beam are determined.

METAL CLUSTERS IN AN ARC PLASMA

Let us consider a plasma in the positive column of a high-pressure arc discharge with a small addition of a refractory metal. This metal is introduced into the plasma in the form of the compound MX_k, where M is the metal atom, X is a halogen atom, and k is the number of halogen atoms in the molecule. The refractory metal exists in the form of these molecules in the cold region of the discharge near the walls, while clusters form in the intermediate region of the discharge. In the central hot region of the discharge the clusters decay, and the refractory metal is found there in the form of atomic vapor.

The equilibrium between the metal atoms and their gaseous component is determined by the processes

$$M + kX \leftrightarrow MX_k . \tag{1}$$

These processes compete with equilibrium processes with the participation of clusters:

$$M_{n-1} + M \leftrightarrow M_n . \tag{2}$$

Let $k\varepsilon_X$ be the total binding energy of the atoms in the molecule MX_k, so that ε_X is the average binding energy in a molecule per halogen atom, and ε_M is the binding energy per atom of the macroscopic metal. We have the following criterion for the existence of the molecules MX_k at low temperatures:

$$\varepsilon_M < k\varepsilon_X \tag{3}$$

Next, on the basis of the equilibrium (1) we have the following relation between the densities of atoms and molecules as a function of temperature T:

$$\frac{[M][X]^k}{[MX_k]} \sim N_0^k \exp\left(-\frac{k\varepsilon_X}{T}\right), \tag{4}$$

where $[A]$ is the density of the atomic particles A and the density N_0 is of the order of an atomic quantity. Similarly, the relation

$$\frac{[\mathrm{M}][\mathrm{M}_{n-1}]}{[\mathrm{M}_n]} \sim N_0 \exp\left(-\frac{\varepsilon_\mathrm{M}}{T}\right) \tag{5}$$

follows from the equilibrium of clusters and atomic vapor. Another version of this equilibrium leads to the relation

$$\frac{[\mathrm{M}][\mathrm{M}_{n-1}]}{[\mathrm{M}_n]} \sim N_{\mathrm{sat}}(T), \tag{6}$$

where $N_{\mathrm{sat}}(T) \sim N_0 \exp(-\varepsilon_\mathrm{M}/T)$ is the density of atoms in the saturated metal vapor at a fixed temperature.

The characteristic temperature T_1 at which these molecules decay into atoms ($[\mathrm{M}] \sim [\mathrm{MX}_k]$) can be estimated from the chemical equilibrium of the molecules MX_k, while the characteristic temperature T_2 at which the clusters transform into atoms can be found from the chemical equilibrium of the clusters. These estimates have the form

$$T_1 = \frac{\varepsilon_\mathrm{X}}{\ln(N_0/[\mathrm{X}])} \quad \text{and} \quad T_2 = \frac{\varepsilon_\mathrm{M}}{\ln(N_0/[\mathrm{M}])}. \tag{7}$$

Here $[\mathrm{X}]$ is the total density of free and bound halogen atoms, while $[\mathrm{M}]$ is the total density of the metal atoms. Evidently, clusters exist in the temperature range

$$T_1 < T < T_2. \tag{8}$$

Hence we find that in the presence of excess halogen atoms $[\mathrm{X}] \gg [\mathrm{M}]$ clusters can exist in the given system if

$$\varepsilon_\mathrm{X} < \varepsilon_\mathrm{M}. \tag{9}$$

Combining the criteria (3) and (9) we obtain a condition under which clusters can exist in the intermediate region of the discharge, while in the cold region the refractory metal is bound in the molecules MX_k:

$$\varepsilon_\mathrm{X} < \varepsilon_\mathrm{M} < k\varepsilon_\mathrm{X}. \tag{10}$$

It is obvious that because of their stability the molecules MX_k have a closed structure. Therefore, in practice, these can be molecules of the type MX_4 or MX_6. The table contains data for such an equilibrium for some molecules of this type. The values of the parameters ε_X and ε_M in the table are taken from handbooks,[9–11] The data in the table refer to a total density of free and bound metal atoms $[\mathrm{M}] = 10^{15}$ cm^{-3} and a total density of halogen atoms $[\mathrm{X}] = 10^{16}$ cm^{-3}. The parameters ε_M and ε_X were determined on the basis of the Gibbs potentials for vapor and compounds.

Expression (7) were used to find the temperatures T_1 and T_2, and the temperature T_3 follows from the relation (6) $[\mathrm{M}] = N_{\mathrm{sat}}(T_3)$. Evidently, the temperatures T_2 and T_3 must be identical, so that the degree to which they are identical attests to the accuracy of the data. We also note that the binding energy of atoms in a cluster is somewhat lower than in a macroscopic system. For example, in the case of tungsten clusters with $n = 1000$ atoms the difference of the binding energies of an atom for a macroscopic system and a cluster is 0.3 eV, which decreases the temperature T_2 by approximately 100 K compared with the tabulated data. It follows from the data in the table and the criteria

obtained that clusters do not form when the molecules TiF_4, ZrF_4, Cl_4, and UF_6 are introduced into an arc discharge, while for other examples in the table clusters do exist in the indicated range of discharge temperatures.

CLUSTER GROWTH AND TRANSPORT IN AN ARC PLASMA

Let us study the kinetics of cluster growth in an arc plasma. It is determined by the processes (2) leading to the attachment of atoms to a cluster and vaporization of clusters as well as by diffusion transport of clusters. For large clusters in the cluster growth zone vaporization is negligible and the balance equation has the form

$$\frac{dn}{dt} = Nk_n, \tag{11}$$

where n is the number of atoms in a cluster, N is the total density of free and bound atoms of the metal, and k_n is the rate constant of the attachment of atoms to a cluster, which for a large cluster, assuming its surface to be similar to that of a macroscopic drop, is

$$k_n = k_0(T)\xi n^{2/3}. \tag{12}$$

Here ξ is the probability of attachment of an atom to the surface of a cluster as a result of atom–cluster contact. In what follows we assume $\xi = 1$, as is the case for a tungsten surface; the rate constant $k_0 = \sqrt{8T/\pi m}\,\pi r_W^2$, where m is the mass of a metal atom and r_W is the Wigner–Seitz radius. Table I contains values of k_0 for the average temperature of the cluster zone. The equation (11) makes it possible to estimate the characteristic cluster size in the cluster region. We shall assume that the displacements of the clusters are determined by their diffusion in the buffer gas, so that the relation between the distance x to which a cluster is displaced in time t is given by the relation $x^2 = 2D_n t$, where D_n is the diffusion coefficient of clusters in the buffer gas. Since the cross section of a cluster varies with the cluster size as $\sim n^{2/3}$, in the liquid-drop model of a cluster $D_n = D_0 n^{-2/3}$, where the diameter D_0 is independent of the cluster size. Then Eq. (11) assumes the form

$$\frac{dn}{dx^2} = \frac{Nk_0}{D_0} n^{4/3}. \tag{13}$$

The solution of this equation for displacements of a cluster whose size varies during the growth process from the initial value n_0 up to the running value n has the form

$$x^2 = \frac{3D_0}{Nk_0}(n_0^{-1/3} - n^{-1/3}). \tag{14}$$

Hence it follows that if the initial distance x_0 of a cluster from the boundary of the cluster region, where cluster formation is thermodynamically favored, satisfies

$$x_0 > \sqrt{\frac{3D_0}{Nk_0 n_0^{1/3}}}, \tag{15}$$

then the growing cluster cannot leave the cluster region, since as the cluster increases in size, the rate of displacements drops rapidly.

TABLE I. Parameters of refractory-metal compounds.

Compound	ε_X, eV	ε_M, eV	T_1, 10^3 K	T_2, 10^3 K	T_3, 10^3 K	k_0, 10^{-11} cm^3/s	n_*
TiF$_4$	6.1	4.9	3.5	2.5	2.3	-	-
TiCl$_4$	4.7	4.9	2.7	2.5	2.3	-	-
TiBr$_4$	3.8	4.9	2.2	2.5	2.3	7.8	$1.8 \cdot 10^4$
ZrF$_4$	6.7	6.1	3.8	3.1	3.1	-	-
ZrCl$_4$	5.1	6.1	2.9	3.1	3.1	2.6	650
ZrBr$_4$	4.4	6.1	2.5	3.1	3.1	2.5	580
MoF$_6$	4.4	6.4	2.5	3.2	3.2	2.5	580
WF$_6$	5.4	8.5	3.0	4.3	4.0	2.0	300
WCl$_6$	3.4	8.5	1.9	4.3	4.0	1.8	220
WBr$_6$	3.0	8.5	1.7	4.3	4.0	1.8	220
IrF$_6$	1.9	6.5	1.1	3.2	3.1	1.5	120
UF$_6$	5.2	4.9	2.9	2.5	2.8	-	-

Thus, the processes considered above lead to a cluster instability, as a result of which metal collects in the growing clusters located in the intermediate region of the arc discharge, while the density of the metal vapor and the molecules present in it decreases with time. Since the concentration of the metal in the buffer gas is very low, processes in which atoms and clusters of the metal participate have no effect on the properties of the discharge. For this reason, the characteristic cluster size is determined only by the life-time of the plasma or the residence time of the metal in the plasma. If the system under study is used to generate a cluster beam, it is convenient to extract plasma from the region of the discharge where clusters accumulate. The simplest method is to allow out of this region of the discharge through an opening. Since the clusters are charged, a beam of clusters is easily separated from the plasma by an electric field. This beam can then be controlled by external fields and can be used for deposition on a substrate by the standard method.[7,8]

The characteristic cluster size in a beam obtained in this manner depends on the growth time of the clusters in the gas-discharge plasma. Specifically, the table contains for the examples studied the characteristic number n_* of atoms in a cluster for the case that the density of free atoms of the metal in the plasma is 10^{15} cm^{-3} and the residence time of the clusters in the plasma for 10^{-3} s. It should be noted that the characteristic cluster size depends sharply on these parameters. This means that the parameters of the cluster beam are highly sensitive to the parameters of the cluster plasma. In addition, we note that the charge of the clusters decreases as the plasma cools, and at low temperatures the clusters become negatively charged. This transition from positive to negative charge of the clusters depends on the electron density. Specifically, for tungsten clusters with $n = 1000$ atoms the transition occurs at temperature $T = 2600$ K if the electron density is 10^{12} cm^{-3}, at temperature $T = 2900$ K if the electron density is 10^{13} cm^{-3}, and at temperature $T = 3290$ K if electron density is 10^{14} cm^{-3}. The cluster charge affects the character of the control of the cluster beam.

In summary, as a result of instability in the cluster region of a high-pressure arc discharge, the state of the system becomes unstable and the clusters grow with time, picking up most of the mass of the refractory metal. This can be the basis for generation

of intense cluster beams. According to the analysis performed above, the characteristic cluster size is sensitive to the density of free and bound metal atoms as well as to the residence time of the metal atoms in the plasma and to the parameters of the plasma flow.

[1] A. V. Eletskiǐ and B. M. Smirnov, Usp. Fiz. Nauk **166**, 1197 (1996).
[2] B. Weber and R. Scholl, J. Illum. Eng. Soc. Summer, 1992, p. 93.
[3] R. Scholl and B. Weber, in *Physics and Chemistry of Finite Systems: From Clusters to Crystals*, Vol. II, edited by P. Jena, S. N. Khana and B. K. Rao, Kluwer Academic Press, Dodrecht, 1992.
[4] B. Weber and R. Scholl, J. Appl. Phys. **74**, 607 (1993).
[5] R. Scholl and G. Natour, in *Phenomena in Ionized Gases*, edited by K. H. Becker, W. E. Carr, and E. E. Kunhardt, AIP Press, Woodbury, 1996.
[6] B. M. Smirnov, Usp. Fiz. Nauk **167**, 1117 (1997).
[7] O. F. Hagena, Z. Phys. D **17**, 157 (1990); *ibid.* **20**, 425 (1991).
[8] T. Takagi, *Cluster Beam Deposition and Epitaxy,* Parker Ridge, Noyes Publications, New York, 1988.
[9] *JANAF Thermochemical Tables*, PB-168370, Clearinghouse, US Department of Commerce, National Bureau of Standards, Washington, D.C., August 1965.
[10] M. Kh. Karanet'yants and M. L. Karapet'yants, *Fundamental Thermodynamic Constants of Inorganic and Organic Substances*, Khimiya, Moscow, 1968.
[11] D. R. Lide (Ed.), *Handbook of Chemistry and Physics*, CRC Press, London, 1998-1999.

Translated by M. E. Alferieff

Quadrupolar forces and aggregation of nematic droplets

I. E. Dzyaloshinskii

Deptartment of Physics, University of California, Irvine, CA 92612, USA

E. I. Kats

Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117940 Moscow, Russia;
Laboratoire de Spectrometrie Physique, Univ. Joseph-Fourier Grenoble 1, France

J. Lajzerowicz

Laboratoire de Spectrométrie Physique, Université Joseph Fourier, Grenoble 1, France

The electrostatic quadrupolar interaction between spherical nematic droplets in an isotropic (and nonconducting) liquid is calculated. It is found to have an anisotropic form $U_q \propto 1/R^5$, where R is the distance between droplets, with repulsion for droplets having parallel orientation of the quadrupole moments and attraction at oblique angles around the orthogonal orientation. In an external magnetic field aligning the orientations of the quadrupole moments, a competition of the quadrupolar repulsion and van der Waals attraction ($U_{\mathrm{vdW}} \propto 1/R^6$) leads to a specific spatial organization of droplets which is in fact often reported in experimental observations (see the monograph by P. Drzaic, *Liquid Crystal Dispersions*, World Scientific, Singapore (1995) and references cited therein). © *1998 American Institute of Physics.*
[S0021-3640(98)00622-7]

PACS numbers: 61.30.−v, 41.20.Cv, 47.55.Dz

1. A nematic liquid crystal is uniform in its ground state, which, however, is very rarely attained in practice. The structure and properties of suspensions (in particular, nematic droplets in an isotropic liquid matrix) pose a number of theoretical and experimental questions which are interesting both from the standpoint of practical applications of liquid crystals and also because phenomena of this kind are of fundamental interest in their own right.

There is one rather evident consequence of nematic ordering that nevertheless, as far as we know, has never been considered before. As is well known,[2] the nematic order parameter is a second-order tensor and has all the properties of the quadrupole moment of a charge distribution. Therefore generally nematics are quadrupolar ferroelectrics (and their isotropic phase is accordingly a quadrupolar paraelectric in which the nematic fluctuations enhance the quadrupolar susceptibility). In spite of this fairly strong statement, in the general case for large (ideally infinite) systems it does not lead to any directly observable predictions, primarily because quadrupolar forces are not very long-

ranged. They are proportional to $1/R^5$, and for large systems this contribution to the energy can be represented (in terms of a Fourier expansion) as being proportional to q^2 (actually to a certain anisotropic combination of $q_i q_j$), and therefore the quadrupolar interactions in principle give some corrections to the bare contributions to the total free energy of the system which scale as q^2 (e.g., compressibility or orientational deformations). From the standpoint of translational symmetry, however, nematics are equivalent to isotropic liquids and must therefore have isotropic ($\propto q^2$) compressibility. Thus the anisotropic quadrupolar contribution to the compressibility in nematics should by definition be zero. The only consequence of quadrupolar forces for bulk nematics is a renormalization of the Frank elastic moduli, which can be represented in the form $K_{ii} = K_{ii}^0 + K_{ii}^Q$, where K_{ii}^0 is the bare value related to forces more short-ranged than the quadrupolar forces, e.g., steric forces, and K_{ii}^Q is the contribution from the quadrupolar interactions. It is clear that such a separation has no relevant physical meaning, and we will not discuss it further in this paper.

The above statement does not necessarily hold for small (mesoscopic) systems, where quadrupolar forces can lead to fairly pronounced phenomena. For a nematic droplet the director distribution field inside the droplet depends on the strength of surface anchoring, the elastic constants, and the droplet size. In the next section some particular configurations having quadrupolar symmetry of the director field will be presented. For such configurations the total orientational quadrupole moment of a droplet is nonzero. On the other hand, in the general case (when the chemical bonds in the material are not purely covalent, and actually they never are 100% covalent — see, e.g., Ref. 3 or the more recent monograph[4]) the droplets should also have an electrostatic quadrupole moment (recall that nematic liquid crystals are quadrupolar ferroelectrics). Therefore an orientational quadrupole moment of the droplet will inevitably lead to an electrostatic quadrupole moment. The latter can create an observable (and not extremely small) electric field around such a droplet (for a droplet radius of the order of 10^{-4} cm it can be as large as 1 V/cm; see the estimates below).

Nevertheless, even for mesoscopic objects the director distribution inside a droplet (and the shape of the droplet itself) is determined mainly by its surface tension and anchoring energy (and, of course, the nematic elastic moduli) and not by the electrostatic quadrupolar energy. Indeed, the electric field created by a quadrupole moment at the droplet surface scales as $E_Q \propto (D_0/L^4)$, where L is the droplet radius and D_0 is the total quadrupole moment, which can be estimated as $D_0 = d_0 N$, where d_0 is the molecular quadrupole moment and N is the number of molecules in the droplet, i.e., $N = (4\pi L^3/3a^2 l)$. It is customarily believed that for most thermotropic nematics the molecules can be represented as hard rods having a well-defined length l and diameter a. Therefore the quadrupolar contribution to the energy is

$$F_{sQ} \simeq \int E^2 d^3 r \simeq \frac{d_0^2}{a^4 l^2} L. \tag{1}$$

This should be compared to the conventional anchoring and surface energy of the nematic droplet, $F_s \simeq W L^2$, where $W \propto T/a^2$ if it is a typical surface energy or $\propto T/l^2$ for the anchoring energy. Thus in any case for

$$L > \frac{d_0^2}{a^4 l^2 W} \simeq l_b \tag{2}$$

(where the so-called Bjerrum length l_b (see, e.g., Ref. 5) is usually of the order of 10 Å, and therefore the condition (2) is always satisfied for any real nematic droplet with $L \simeq 10^5 - 10^4$ Å), and the shape (and director distribution) of the nematic droplet is determined mainly by the conventional surface energy.

However, let us stress once again that the quadrupolar forces (though they give only small corrections to the thermodynamic properties) are not negligible. For example, using the rough estimates given above, one can find that for a droplet with $L \simeq 10^{-4}$ cm the quadrupolar electric field around it will be about $E_Q \simeq 1$ V/cm.

2. From what we have said above it is clear that under the condition (2) the director distribution in a nematic droplet can be found in the standard way[2] by minimization of the Frank elastic energy with suitable boundary conditions. In fact, a simple qualitative analysis of possible director configurations shows that for a spherical droplet with any boundary condition (the only exception being an infinitely strong radial anchoring), the director distributions always have quadrupolar symmetry. We might mention, for example, bipolar structures with two surface point defects (so-called boojums), which can be realized for tangential boundary conditions, or more-sophisticated (but also possessing quadrupolar symmetry) structures which arise in the case of tilted orientation of the director on the droplet surface and which correspond to two boojums, one hedgehog, and one disclination ring (see Ref. 6). All these configurations can be characterized by a quadrupolar preferred direction $\mathbf{m} \equiv -\mathbf{m}$ (e.g., for the bipolar structure $\mathbf{m}$ is a unit vector along the direction between two poles). Thus the quadrupole moment of such a configuration can be represented in the form

$$D_{ik} = \tilde{D}_0 \left(m_i m_k - \frac{1}{3} \delta_{ik} \right), \tag{3}$$

and the only difference between different quadrupolar configurations is the numerical coefficient (of the order of 1/10) which relates $\tilde{D}_0$ and D_0. We henceforth neglect this difference, or, more accurately, consider this numerical factor to be included in the definition of $\tilde{D}_0$.

The quadrupole–quadrupole pair interaction between droplets having quadrupole moments $\hat{D}^{(1)}$ and $\hat{D}^{(2)}$ for distances R larger than the size of each droplet can be found easily for the case when the droplets have no charge and no dipole moment:[7]

$$F_{\text{int}} = \frac{1}{12} D_{\alpha\beta}^{(2)} \frac{\partial}{\partial x_\alpha} \frac{\partial}{\partial x_\beta} \left[D_{\gamma\delta}^{(1)} \frac{n_\gamma n_\delta}{R^3} \right], \tag{4}$$

where $\mathbf{n}$ is a unit vector along $\mathbf{R}$. Using (3) and performing the calculations, we find

$$F_{\text{int}} = \frac{\tilde{D}_0^{(1)} \tilde{D}_0^{(2)}}{12 R^5} \left[(\mathbf{m}_1 \cdot \mathbf{m}_2)^2 - 20 (\mathbf{n} \cdot \mathbf{m}_1)(\mathbf{n} \cdot \mathbf{m}_2)(\mathbf{m}_1 \cdot \mathbf{m}_2) - 5[(\mathbf{n} \cdot \mathbf{m}_1)^2 + (\mathbf{n} \cdot \mathbf{m}_2)^2] \right.$$

$$\left. + 35 (\mathbf{m}_1 \cdot \mathbf{n})^2 (\mathbf{n} \cdot \mathbf{m}_2)^2 + \frac{4}{3} \right]. \tag{5}$$

In the general case F_{int} is a function of three angular variables: θ_1, θ_2, and $\varphi_1 - \varphi_2$ (where θ_i and φ_i are, respectively, the polar and azimuthal angles between $\mathbf{m}_i$ and $\mathbf{n}$).

Although the complete investigation of this function can be done only numerically, some particular configurations of relevance to us can rather easily be calculated analytically. Namely, one can show that attraction takes place only for nearly orthogonal orientations[a] of $\mathbf{m}_1$ and $\mathbf{m}_2$, while for nearly parallel vectors $\mathbf{m}_i$ we have always repulsion. Expression (5) has a zero average over a sphere, as it should for quadrupolar symmetry.

Obviously $F_{\text{int}}(R, \theta_1, \theta_2, \varphi_1 - \varphi_2)$ is a noncentral potential, and the force on the nematic droplet has both radial and angular (torque) components. Therefore when we have many droplets interacting with each other in many different directions, and also as a result of Brownian motion, these droplets will always find a way towards the global angular minimum. This implies that the droplets would tend to organize themselves into chains with alternating orthogonal orientations of the quadrupole moments of neighboring droplets. One can consider organization of this type as an analogy of ferroelectric domains for quadrupolar ferroelectrics.

It is easy to estimate the characteristic time for chain formation controlled by the Stokes friction force $6\pi\eta L u$, where η is the viscosity of the liquid and u is the velocity of a droplet. The interdroplet distance R is proportional to $\phi^{-1/3}L$, where ϕ is the droplet concentration. Therefore using expression (5) for quadrupolar attraction and equating the frictional and quadrupolar forces, one can estimate the characteristic time as

$$\tau \simeq \frac{\eta a^4 l^2}{d_0^2} L^2 \phi^{-7/3}. \tag{6}$$

Here (as well as later on) we ignore the entropy of mixing of the droplets, treating them as macroscopic objects, and, of course, the estimate (6) is valid only for small ϕ, when $R \gg L$.

Interesting phenomena can appear when we apply an external magnetic field to the system. Recall that the droplets are not only quadrupolar ferroelectrics but also conventional nematics having a diamagnetic anisotropy χ_a. Therefore a sufficiently strong magnetic field can provide alignment of the droplet quadrupole moments. The necessary condition for this alignment is

$$\chi_a^2 H^2 L^3 > \frac{\tilde{D}^2}{R^5}, \tag{7}$$

i.e., the magnetic energy must overcome the quadrupolar attraction. From (7) we find the critical field

$$H_c = \frac{d_0}{\sqrt{\chi_a} a^2 l} \frac{1}{L} \tag{8}$$

(we have assumed here that $R \simeq L$). Natural estimates of the parameters appearing in (8) give $H_c \simeq 10^3 - 10^4$ G.

However to avoid director deformations inside a droplet (and the consequent reduction of the droplet quadrupole moments) under the influence of this field, it must be smaller than the critical field for the Fréedericksz transition[2,9]

$$H_F = \sqrt{\frac{K}{\chi_a}} \frac{1}{L}. \tag{9}$$

Comparing (8) and (9), we get the condition

$$d_0 < \sqrt{K} a^2 l. \tag{10}$$

Estimating K as T/l, we conclude that the condition (10) is equivalent to the condition

$$l_b < l\left(\frac{l}{a}\right),$$

which can be satisfied for real nematics.

Under these conditions (namely in an external magnetic field $H_c < H < H_F$) the quadrupolar forces give repulsion between droplets (as can be easily seen from (5)), and the chain structure will therefore be broken. To find the structure which will arise as a result of this quadrupolar repulsion ($\propto 1/R^5$) one has to take into account the others forces acting in such a system. The most relevant among these are the van der Waals forces. These forces lead primarily to a certain contribution to the internal energy of each droplet. This internal contribution can be calculated using the general method elaborated by Dzyaloshinskii, Lifshitz, and Pitaevskii (see, e.g., Ref. 10; some features of this contribution which are specific to nematics and which give, in particular, a renormalization of the Frank moduli, have been studied in Ref. 11).

For us it is more relevant that the van der Waals forces lead to attraction between droplets. This attraction is anisotropic; in any case, however, the average over a sphere is nonzero. Thus we have a rather unusual situation: a system of particles with, say, ''long-range'' ($1/R^5$) repulsion and ''short-range'' ($1/R^6$) attraction.

The detailed calculation of all possible configurations for such systems would address a very complicated kinetic (and statistical) problem which is beyond the scope of our paper. A more modest aim we have is to describe qualitatively the consequences of quadrupolar forces. At this level it is clear that in the presence of both forces (the quadrupolar and van der Waals) the droplet distribution can be characterized by a certain specific length scale Λ. The above discussion of the forces assumed $1/R^5$ or $1/R^6$ laws for the interactions between isolated molecules. To find the total interactions between two mesoscopic spheres one has to sum of all individual interactions. This reduces to evaluation of the following integral:

$$I = \int d^3 r_1 \int d^3 r_2 \frac{1}{S^\alpha}, \tag{11}$$

where

$$S^2 = [(R + 2L - r_2 \cos \theta_2 - r_1 \cos \theta_1)^2 + (r_2 \sin \theta_2 - r_1 \sin \theta_1)^2], \tag{12}$$

and the exponent $\alpha = 3$ for the van der Waals interactions and $\alpha = 5/2$ for the quadrupolar interactions. In the simplest case $R \ll L$ it follows from (11)–(12) that

$$I_{\text{vdW}} \propto \frac{L}{R}, \quad I_Q \propto \ln \frac{L}{R}.$$

Now we are in position to estimate the characteristic cluster scale as the scale at which the van der Waals attraction is of the same order as the quadrupolar repulsion. Putting together all dimensional factors, we get

$$\Lambda \simeq \frac{h \nu \alpha_0^2 a^4 l^2}{d_0^2} \ln \frac{L}{l},$$

where α_0 is the electronic polarizability and ν is a characteristic frequency that can be identified with the first ionization potential of the molecules, which usually falls in the ultraviolet region (the meaning of all the others parameters has been explained above).

3. We calculated the electrostatic quadrupolar interaction between spherical nematic droplets in an isotropic (and nonconducting) liquid. It has an anisotropic form $U_q \propto 1/R^5$, where R is the distance between droplets, with repulsion for droplets having parallel orientation of the quadrupole moments and with attraction at oblique angles around the orthogonal orientation. In an external magnetic field ordering the orientations of the quadrupole moments, a competition of the quadrupolar repulsion and van der Waals attraction ($U_{\text{vdW}} \propto 1/R^6$) leads to a specific spatial aggregation of droplets that has in fact been reported often in experimental observations (see, e.g., Refs. 1, 8, 9, 12, 4, and 13). Results analogous to those given above should be applicable as well to colloidal suspensions of quadrupolar polarizable particles in a less-polarizable nonconducting fluid.

Obviously the above description is rather simplistic and does not include any un-welcome effects, namely:

i) Charge screening. To neglect this effect we assumed that the isotropic liquid is ideally nonconducting.

ii) Depletion.[14,13] It leads only to a short-range attraction between spheres with a characteristic scale of the order of l.

iii) Coarsening and coalescence. Neglecting these processes is a good approximation at a sufficiently small droplet concentration.

iv) Flexoelectric polarization. Flexoelectricity in nematic liquid crystals is a phenomenon that has been studied in many publications for over a decade (see, e.g., Refs. 2 and 8). It is known[2] that deformations of the director field in nematics can create a polarization

$$\mathbf{P}_f = e_1 (\mathbf{n} \nabla \cdot \mathbf{n}) + e_3 (\nabla \times \mathbf{n}) \times \mathbf{n},$$

which involves two coefficients with dimensions of electric potential. In the case of molecules which are very asymmetric in shape and carry a strong electric dipole moment μ_d the flexoelectric coefficients might reach values of order μ_d / l^2. In all other cases (and in particular in the case under consideration, when the molecules do not have a permanent dipole moment) they will be smaller. Rough estimates of these coefficients based on calculations of the fraction of the molecules which achieve the necessary ordering of their dipoles to ensure the maximum packing density give[8]

$$e \propto N^{1/3}/T,$$

where N is the number of molecules per unit volume. Thus the corresponding electric field (proportional to $N^{1/3}$) will be small in comparison to the quadrupolar field, which is proportional to N. Note also the following difference between the two kinds of polarizations. The quadrupolar polarization (at least in principle) can be nonzero even for **n** = const, while the flexoelectric polarization is proportional to gradients of **n**.

[a] The angular interval in which attraction takes place can be characterized by a certain angle between **n** and one of the vectors **m**, which is varied from $\arccos\sqrt{(4/15)}$ to $\arccos\sqrt{(2/3)}$, and the maximum is achieved for precisely orthogonal orientations.

[1] P. Drzaic, *Liquid Crystal Dispersions*, World Scientific, Singapore (1995).

[2] P. G. de Gennes, *The Physics of Liquid Crystals*, Clarendon Press, Oxford (1974).

[3] A. I. Kitaigorodsky, *Molecular Crystals and Molecules*, Academic Press, New York (1973).

[4] D. Myers, *Surfaces, Interfaces, and Colloids*, VCH Publishers, New York (1991).

[5] J. N. Israelashvili, *Intermolecular and Surface Forces*, Academic Press, Orlando (1985).

[6] G. E. Volovik and O. D. Lavrentovich, Zh. Éksp. Teor. Fiz. **85**, 1997 (1983) [Sov. Phys. JETP **58**, 1159 (1983)].

[7] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Pergamon Press, New York (1974).

[8] S. Chandrasekhar, *Liquid Crystals*, Cambridge Univ. Press, New York (1977).

[9] E. Dubois-Violette and O. Parodi, J. Phys. (Paris), Colloq. **30**, C4-57 (1969).

[10] I. E. Dzyaloshinskii, E. M. Lifshitz, and L. P. Pitaevskii, Adv. Phys. **10**, 165 (1961).

[11] I. E. Dzyaloshinskii, S. G. Dmitriev, and E. I. Kats, Zh. Éksp. Teor. Fiz. **68**, 2335 (1975) [Sov. Phys. JETP **41**, 1167 (1975).

[12] W. B. Russel, D. A. Saville, and W. R. Schowalter, *Colloidal Dispersion*, Cambridge University Press, Cambridge (1989).

[13] P. Poulin, V. Cabuil, and D. A. Weitz, Phys. Rev. Lett. **79**, 4862 (1997).

[14] Y. Mao, M. E. Cates, and H. N. W. Lekkerkerker, Phys. Rev. Lett. **75**, 4548 (1995).

On the nature of the oscillations of cyclotron absorption in InAs/GaSb quantum wells

S. D. Suchalkin[a)] and Yu. B. Vasil'ev

A. F. Ioffe Physicotechnical Institute, Russian Academy of Science, 194021 St. Petersburg, Russia; Max-Planck-Institut für Festkörperforschung, 70569 Stuttgart, Germany

K. von Klitzing

Max-Planck-Institut für Festkörperforschung, 70569 Stuttgart, Germany

V. N. Golovach, G. G. Zegrya, S. V. Ivanov, P. S. Kop'ev, and B. Ya. Mel'tser

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

The mechanism of oscillations of the half-width and intensity of the cyclotron resonance (CR) line of electrons in a semimetal quantum well based on an InAs/AlSb/GaSb heterostructure is investigated experimentally and theoretically. It is shown that the oscillations of the CR spectrum are due to mixing of states of the spatially separated two-dimensional electrons and holes. © *1998 American Institute of Physics.* [S0021-3640(98)00722-1]

PACS numbers: 76.40.+b, 73.61.Ey

The GaSb valence-band top in an InAs/AlSb/GaSb semiconductor heterostructure lies approximately 150 meV above the InAs conduction-band bottom. The uniqueness of such structures lies in the fact that spatially separated two-dimensional electrons and holes with overlapping energy spectra can coexist in them.[1,2] The investigation of cyclotron resonances (CRs) in such structures has shown that the absorption spectrum possesses a number of characteristic features, specifically, a strong oscillatory magnetic-field dependence of the half-width and intensity of the CR line is observed. In the first publications such oscillations were interpreted as being a manifestation of the dependence of the efficiency of the screening of a chaotic scattering potential on the filling factor of the Landau levels.[3] However, recent investigations of CRs in InAs/Al$_x$Ga$_{1-x}$Sb quantum wells[4] have shown that the oscillations are observed only in semimetal-type samples ($x < 0.3$), i.e., where electrons and holes are present. This fact enabled the authors of Ref. 4 to infer that the observed features in the CR spectra are due to the Coulomb interaction between electrons and holes. In theoretical work[5] the characteristic behavior of the shape of the cyclotron resonance line in InAs/GaSb quantum wells was attributed to the characteristic features of the band structure in the energy range corresponding to overlapping of the InAs conduction band and the GaSb valence band.

Our objective in the present work is to investigate experimentally and theoretically
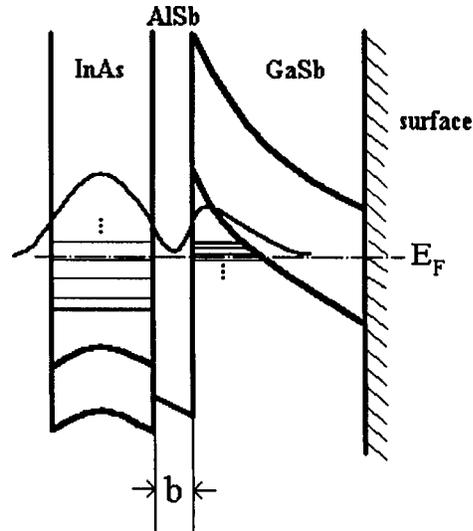
FIG. 1. Band diagram of the experimental structure. The wave function of a hybrid electron–hole state is shown schematically.

the mechanisms leading to oscillations of the half-width and intensity of the CR line in an InAs/AlSb/GaSb heterostructure.

Samples grown by molecular beam epitaxy were used in the experiment. The samples consisted of 200 Å wide InAs quantum wells, separated from the GaSb matrix by AlSb barriers. A series of structures with barrier thicknesses 0, 6, 20, 100, and 300 Å was investigated. Charge carriers in InAs quantum wells accumulate without additional doping. The main sources of the carriers are surface states with energy close to the center of the gap E_g in GaSb and states in the valence band of GaSb. The transition of electrons from states into the quantum well results in the appearance of a self-consistent quantum well for holes in GaSb (Fig. 1b). The CR spectra of the electrons in InAs quantum wells were investigated in magnetic fields ranging from 2 to 13 T at temperature $T = 2.2$ K. The measurements were performed with a Bruker IFS-113v Fourier spectrometer in a fixed magnetic field. The probe radiation was detected with a Si bolometer. All spectra were normalized to the spectrum obtained without a magnetic field. The magnetotransport measurements were performed in order to estimate the electron density, which was varied by illuminating the samples with light from a red LED.

The main experimental problem was to study the transformation of the CR spectra with increasing width of the AlSb barrier separating electrons and holes. The presence of holes in all experimental samples was confirmed by observing the CR line corresponding to an effective cyclotron mass $m^* \approx 0.3 m_0$. The CR of electrons in a sample with 6 Å barriers (Fig. 2a) shows strong oscillations of the half-width and intensity of the absorption peak, similar to the results obtained in Ref. 3 on InAs/GaSb quantum wells without a barrier. The oscillations are periodic in $1/B$, and the maximum half-widths of the CR line correspond to the minimum amplitudes of the line. Such a magnetic-field dependence of the CR line shape was also observed in a sample without an AlSb barrier. This attests to the fact that the appearance of oscillations is not associated with the properties of the
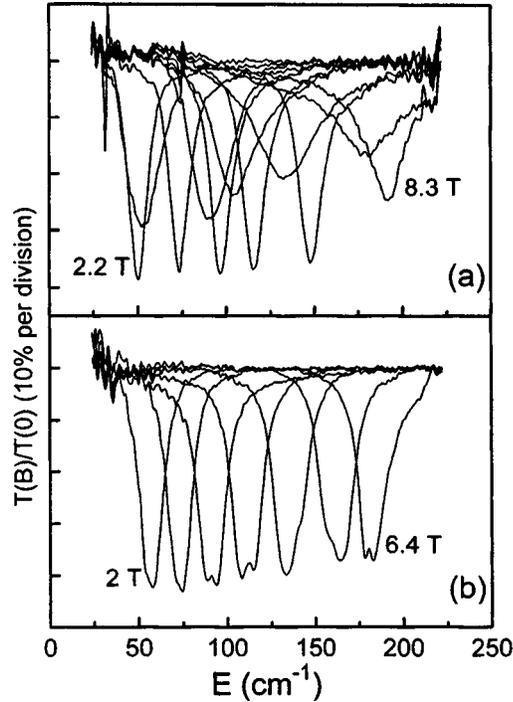
FIG. 2. Absorption spectrum of an InAs/AlSb/GaSb quantum well for different values of the magnetic field. AlSb barrier thicknesses: a) $b = 6$ Å , b) $b = 20$ Å .

interface separating electron and hole layers. As the barrier thickness increases to 20 Å , the cyclotron absorption spectrum as a function of the magnetic field changes radically. Figure 2b shows the CR spectra of a sample with 20 Å barriers. Instead of strong oscillations of the half-width and intensity of the absorption line as a function of the magnetic field, splitting of the CR line is observed in certain magnetic fields. A detailed description of this phenomenon will be presented in subsequent publications. Here we would like to discuss the physical reasons for the behavior of the CR in these structures accompanying a transition from GaSb barriers to AlSb barriers. The basic idea of the experimental part of this work is that this transition is made not by changing the Al fraction in the barrier material, as done in Ref. 4, but by increasing the barrier thickness, starting with the lowest value at which the AlSb unit cell remains (~ 6 Å). The experimental results showed that the oscillations of the half-width and intensity of the CR line vanish as the AlSb barrier thickness increases from 6 to 20 Å .

To analyze the experimental results obtained, it is first necessary to investigate the effect of hole states from GaSb on the spectrum and wave functions of electrons in InAs. For this, we shall employ a four-band Kane model. The wave function of the carriers can be represented as

$$\psi = U|s\rangle + \mathbf{V}|\mathbf{p}\rangle, \tag{1}$$

where $|s\rangle$ and $|\mathbf{p}\rangle = \{|X\rangle, |Y\rangle, |Z\rangle\}$ are s- and $\mathbf{p}$-type Bloch functions with angular momentum 0 and 1, respectively; $U(\mathbf{r})$ and $\mathbf{V}(\mathbf{r}) = \{V_x(\mathbf{r}), V_y(\mathbf{r}), V_z(\mathbf{r})\}$ are smooth enve-

lopes of the Bloch functions, which are spinors. Near the Γ point the system of equations for the envelopes in a spherical approximation in the presence of a magnetic field has the form[6]

$$(E_c - \boldsymbol{\mu} \cdot \mathbf{H} - E)U + \gamma \hat{\mathcal{K}} \cdot \mathbf{V} = 0,$$

$$\gamma \hat{\mathcal{K}} U + \left(E_v - \delta - \boldsymbol{\mu} \cdot \mathbf{H} - (\tilde{\gamma}_1 - 2\tilde{\gamma}_2) \frac{\hbar^2}{2m} \hat{\mathcal{K}}^2 - E \right) \mathbf{V}$$

$$- \tilde{\gamma}_2 \frac{3\hbar^2}{m} \hat{\mathcal{K}}(\hat{\mathcal{K}} \cdot \mathbf{V}) + i\delta[\boldsymbol{\sigma} \times \mathbf{V}] = 0. \tag{2}$$

Here γ is the Kane matrix element, $\tilde{\gamma}_1$ and $\tilde{\gamma}_2 = \tilde{\gamma}_3$ are generalized Luttinger parameters, $\delta = \Delta_{so}/3$, Δ_{so} is the spin–orbit splitting constant, E_c and E_v are, respectively, the energies of the conduction-band bottom and valence-band top ($E_g = E_c - E_v$), m is the mass of a free electron, $\boldsymbol{\sigma} = \{\sigma_x, \sigma_y, \sigma_z\}$ are Pauli matrices, $\boldsymbol{\mu} = -\frac{1}{2} g_0 \mu_B \boldsymbol{\sigma}$ is the magnetic moment of the electron, $\mathbf{H}$ is the magnetic field intensity ($\mathbf{H} = \{0,0,H\}$), μ_B is the Bohr magneton, g_0 is the g factor of a free electron, $\hat{\mathcal{K}} = -i\boldsymbol{\nabla} - (e/c\hbar)\mathbf{A}$ is the generalized wave vector of a particle ($\hat{k} \equiv \hat{\mathcal{K}}_z, \hat{k}_z \equiv -i(\partial/\partial x)$), $\mathbf{A}$ is the vector potential of the field ($\mathbf{A} = \{-yH, 0, 0\}$). The wave function of the state $N\uparrow$ has the form

$$\psi_{N\uparrow} = \begin{pmatrix} \alpha_1 \uparrow + \beta_1 \downarrow \hat{a}^+ \\ \alpha_2 \uparrow \hat{a}^+ + \beta_2 \downarrow (\hat{a}^+)^2 \\ \alpha_3 \uparrow \hat{a} + \beta_3 \downarrow \\ \alpha_4 \uparrow + \beta_4 \downarrow \hat{a}^+ \end{pmatrix} \varphi_{k_x}(x) \chi_N(y) f_k(z), \tag{3}$$

where $\varphi_{k_x}(x) \propto \exp(ik_x x)$, $\chi_N(y)$ are harmonic oscillator functions, and the function $f_k(z)$ satisfies $\hat{k}^2 f_k(z) = k^2 f_k(z)$. In expression (3), for convenience, a transition was made from the basis vectors $|X\rangle, |Y\rangle$ to the basis vectors $|p_-\rangle, |p_+\rangle$, respectively, according to the rules $|p_-\rangle = 2^{-1/2}(|X\rangle - i|Y\rangle)$ and $|p_+\rangle = -2^{-1/2}(|X\rangle + i|Y\rangle)$); in addition, creation and annihilation operators $\hat{a}^+ = -2^{-1/2} a_H (\hat{\mathcal{K}}_x + i\hat{\mathcal{K}}_y)$ and $\hat{a} = -2^{-1/2} a_H (\hat{\mathcal{K}}_x - i\hat{\mathcal{K}}_y)$, where $a_H = \sqrt{\hbar c/|e|H}$ is the magnetic length, are used. The components of the spinors $\alpha_i \uparrow = \binom{\alpha_i}{0}$ and $\beta_i \downarrow = \binom{0}{\beta_i}$ can contain the operator $\hat{k}$. The state $(N+1)\downarrow$ also possesses a wave function of the same general form (3). Therefore mixing of the states $N\uparrow$ and $(N+1)\downarrow$ occurs in the quantum well.

The boundary conditions formulated in Ref. 6 were used to find the spectrum of the carriers in the quantum well.

To analyze the hybridization of the states of the conduction band of InAs and the states of the valence band of GaSb, it is important to establish the selection rules for resonance tunneling of charge carriers through an AlSb barrier. Taking into account the mixing of the states at the heterointerface, we find that only the following matrix elements of resonance tunneling are nonzero:

$$\langle \psi^e_{N\uparrow} | \psi^h_{N\uparrow} \rangle, \quad \langle \psi^e_{N\uparrow} | \psi^h_{(N+1)\downarrow} \rangle,$$

$$\langle \psi^e_{N\downarrow} | \psi^h_{N\downarrow} \rangle, \quad \langle \psi^e_{N\downarrow} | \psi^h_{(N-1)\uparrow} \rangle. \tag{4}$$

The selection rules (4) reflect the fact that the Hamiltonian computes with the operator

$$\hat{j} = \hat{a}^+ \hat{a} + \hat{\ell}_z + \hat{s}, \tag{5}$$

where $\hat{\ell}_z$ is the projection operator of the orbital angular momentum of a particle onto the z axis, and $\hat{s} = \frac{1}{2}\sigma_z$ is the projection operator of the spin of a free electron onto the z axis. Using the form of the wave function (3), we find the eigenvalues of the operator $\hat{j}$

$$\hat{j}\psi_{N_s} = (N+s)\psi_{N_s}, \tag{6}$$

where $s = \pm 1/2$, while the quantum number N is a sum of the eigenvalues of the operators $\hat{\ell}_z$ and $\hat{n} \equiv \hat{a}^+\hat{a}$: $N = n + \ell_z$. For electrons $\ell_z = 0$, and therefore the numbers of the Landau levels are the same as the eigenvalues of the operator $\hat{n}$. For carriers in the valence band in a magnetic field mixing of states of heavy ($\ell_z = \pm 1$), light and spin-split-off ($\ell_z = 0$) holes mix, and for this reason the Landau levels for holes cannot be enumerated with the quantum number n.

Since the operator $\hat{j}$ does not depend on the coordinate z, the quantity $j = N + s$ is an invariant for resonance tunneling of carriers through a barrier. For tunneling with conservation of the orientation of the electron spin, the tunneling matrix element can be accurately represented as

$$\langle \psi^e_{N\uparrow} | \psi^h_{N\uparrow} \rangle \approx \sqrt{\frac{N}{ac}} \frac{k_h}{k_h^2 + \varkappa_e^2} \frac{\gamma}{a_H E_g} e^{-\varkappa_e b} \cos k_e a, \tag{7}$$

where b is the width of the AlSb barrier, a and c are, respectively, the half-widths of the electron and hole quantum wells (for simplicity, we approximated the self-consistent hole quantum well by a square well), k_e and k_h are, respectively, the z components of the electron and hole quasimomenta, and $\varkappa_e^{-1}$ is the characteristic decay length of an electron wave function under the AlSb barrier. The matrix element of the tunneling transition with spin flip has the same exponential dependence on the barrier thickness (see Eq. (7)).

The wave function of carriers in an InAs/GaSb heterostructure is a superposition of electron and hole states. However, the mixing of these states is substantial only when the conditions for resonance tunneling of carriers between electron and hole wells are satisfied (see the selection rules (4)). It is obvious that the admixing of hole states to electron states will have a large effect on the shape of the electronic CR line, since the mobility of the holes is much lower than that of the electrons. The stronger this admixture, the wider the cyclotron absorption line is. Therefore the broadening occurs for values of the magnetic field for which the energies of the Landau levels of the electron and hole systems coupled by the selection rules determined above are the same. In Fig. 3 the arrows mark the computed values of the magnetic field for which resonance tunneling of charge carriers through the AlSb barrier occurs. The solid arrows correspond to tunneling without spin flip, while the broken arrows correspond to tunneling with spin flip. The width of the hole well and the quantity $g_h - g_e$, where g_e and g_h are, respectively, the electron and heavy-hole g factors, were used as adjustable parameters. The degree of hybridization of the ''interacting'' levels, as follows from Eq. (3), depends exponentially on the thickness of the AlSb layer. As the thickness of the AlSb barrier increases from 6 to 20 Å, the probability of electron tunneling through the barrier decreases by an order of magnitude,
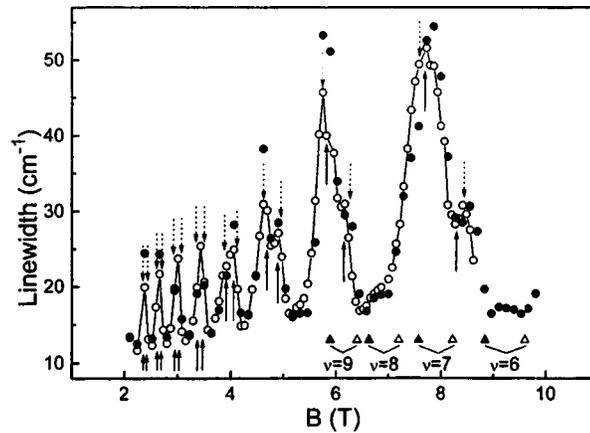
FIG. 3. Half-width of the CR line versus the magnetic field in an InAs/AlSb/GaSb quantum well with 6-Å barriers before (○) and after (●) illumination with an LED. The triangles mark magnetic fields corresponding to definite values of the filling factor before (△) and after (▲) illumination. The arrows mark the computed values of the magnetic field for which the broadening of the CR line is assumed to occur. The solid arrows correspond to resonance tunneling with conservation of spin orientation, while the broken arrows correspond to tunneling with spin flip.

and the oscillations of the half-width and intensity of the CR line practically vanish (Fig. 2). It is obvious that the change in the electron–hole distance by such a small amount cannot qualitatively influence the long-range Coulomb interaction.[7]

It should be noted that the oscillations of the half-width and intensity of the CR line, in contrast to the Shubnikov–de Haas oscillations, are not determined by the filling factors of the Landau levels. Figure 3 shows the magnetic-field dependence of the half-width of the CR line in a sample with a fixed 6-Å barrier before and after illumination with by an LED, which decreased the electron density in the well by approximately 10% (negative frozen photoconductivity effect[8]). The triangles mark the magnetic fields corresponding to integral values of the filling factor before (open triangles) and after (filled triangles) illumination of the sample. One can see that even though a clear shift is observed in the Shubnikov–de Haas oscillations (Fig. 3), the magnetic fields corresponding to the extremal values of the half-width of the electron CR peak remain unchanged. This is explained by the fact that the period of the oscillations of the CR line shape as a function of the reciprocal of the magnetic field in InAs/GaSb ''semimetal'' quantum wells is determined not by the Fermi energy but by the energy between the electron and hole size-quantization subbands, which depends weakly on the carrier density.

[a]e-mail: S.Suchalkin@pop.ioffe.rssi.ru

[1] J. C. Maan, Y. Guildner, J. P. Vieren *et al.*, Solid State Commun. **39**, 683 (1981).
[2] L. S. Kim, H. D. Drew, and H. Munekata, Solid State Commun. **66**, 873 (1988).
[3] D. Heitmann, M. Ziesmann, and L. L. Chang, Phys. Rev. B **34**, 7463 (1986).
[4] J. Kono, B. D. McCombe, J.-P. Cheng *et al.*, Phys. Rev. B **55**, 1617 (1997).
[5] J.-C. Chiang, S.-F. Tsay, Z. M. Chau, and I. Lo, Phys. Rev. Lett. **77**, 2053 (1996).

[6] G. G. Zegrya and A. S. Polkovnikov, Zh. Éksp. Teor. Fiz. **113**, 1491 (1998) [JETP **86**, 815 (1998)].
[7] R. Lassing and E. Gornik, Solid State Commun. **47**, 959 (1983).
[8] C. Nguyen, B. Brar, and H. Kroemer, Appl. Phys. Lett. **60**, 1854 (1992).

Translated by M. E. Alferieff

On the stability of a moving charged helium film

V. B. Shikin[a)]

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

The problem of the stability of a massive, charged helium film is studied with allowance for the possible motion of the film relative to the electron system localized on its surface. Here a peculiar sort of Doppler effect is intertwined with processes that cause the charged helium surface to become unstable, affecting its critical characteristics. © *1998 American Institute of Physics.* [S0021-3640(98)00822-6]

PACS numbers: 67.70.+n

Research on the properties of two-dimensional (2D) charged systems moving relative to a liquid substrate has become increasingly active in recent years. This trend is best displayed in various attempts to investigate the properties of an electron crystal sliding along a liquid substrate (see, e.g., Refs. 1–3). As is well known, a stationary electron lattice deforms a liquid substrate, forming craters under each of the electrons localized on lattice sites. The dynamical properties of an electron crystal on a liquid substrate are largely determined by this deformation.[4] However, if the crystal moves with a definite velocity relative to the liquid surface, there is not enough time for craters to appear and the crystal acquires new properties.

Another interesting problem involving the relative motion between electrons and a liquid substrate arises from ''hydrodynamic'' considerations. It is well known that solitary waves (solitons) can propagate along a neutral channel of finite depth.[5] The appearance of a finite charge density on a liquid surface should seriously modify the properties of the hydrodynamic soliton. More generally, there is the problem of finite-amplitude third-sound on a moving charged helium film, including the effect of the motion of the film on its stability criterion.

Existing attempts to force the electron system to slide as a whole relative to a liquid substrate are ineffective. Here the problem is to use electric forces[1] or Lorentz forces[2] along the helium surface. Of course, these forces displace the electrons relative to the substrate, but sliding is spatially extremely nonuniform and appears only in a threshold manner. As a result, the observed effects are difficult to interpret.

In the present letter, attention is called to a comparatively simple, alternative possibility of obtaining uniform motion of an electron system relative to helium without using perturbing electric forces. The problem is to produce conditions under which the liquid helium film is forced to flow along a solid substrate. The superfluid nature of the helium makes it possible to accomplish this. As a result, the electronic component on a liquid
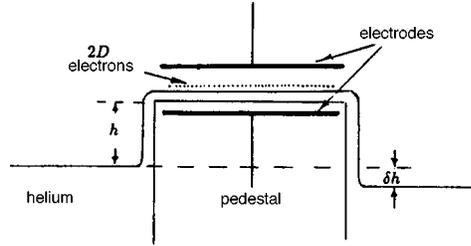
FIG. 1. Diagram of the preparation of a 2D electron system on a moving substrate.

surface of a helium film remains stationary in the laboratory coordinate system. The film, however, flows relative to the solid substrate and therefore relative to the electron system. The stability of a mobile charged helium film is investigated below on the basis of this scenario.

1. Let us consider the cell shown schematically in Fig. 1. The electrons are localized on a special pedestal, which is raised relative to the level of the bulk helium to height h. This makes it possible to regulate the thickness d of the film on the pedestal:[4]

$$d^{-1} = \left(\frac{1}{d_0^3} + \frac{2\pi e^2 n_s^2}{f} \right)^{1/3} \quad \text{and} \quad d_0 = \left(\frac{f}{\rho g h} \right)^{1/3}, \tag{1}$$

where g and ρ are the acceleration of gravity and density of liquid helium, f is the van der Waals constant, and n_s is the equilibrium density of the surface electrons. In addition, a definite differential δh of the heights of the bulk helium to the left and right of the pedestal is provided. This differential allows superfluid flow of helium from left to right or vice versa, depending on the sign of δh, with velocity $u \leqslant u_c$, where u_c is the critical velocity of the superfluid flow along the film and is given by

$$u_c = \frac{2\hbar}{m_4 d} \ln(d/a), \tag{2}$$

where m_4 is the mass of a helium atom and a is the interatomic distance. For $d \sim 10^{-1}$ cm, $u_c \sim 0.62$ cm/s.

The position of the electrons of density n_s on the pedestal is controlled by the fields produced by a flat-plate capacitor that holds the electrons on the liquid substrate. This distribution is, on the average, stationary relative to the rises and therefore it is in motion relative to the liquid film.

2. The dispersion relation of a capillary wave $\exp(iqx - i\omega t)$ on the surface of a moving charged helium film under saturation conditions (the electric field above the film is zero) is

$$\omega_{\pm} = qu \pm \left[\frac{|q|}{\rho} (\rho g^* + \alpha q^2 - 4\pi\sigma^2 |q|) \right]^{1/2}, \tag{3}$$

$$g^* = g + 3f/\rho d^4, \quad q > d,$$

where $\sigma = e n_s$ and α is the surface tension of liquid helium. Including the contribution of van der Waals forces in the definition of g^* makes it possible to use the dispersion relation (3) for quite thin helium films.

To obtain the obvious result (3) (Doppler effect in the problem of ripplon oscillations) it is necessary not only to modify the boundary condition for the hydrodynamic potential ϕ

$$\frac{\partial \phi}{\partial t} = u \frac{d\xi}{dx} + \frac{d\xi}{dt}$$

(ξ is the displacement of the liquid boundary from the equilibrium position) but also to take into account a correction linear in ϕ to the general pressure balance on the mobile boundary as a result of the Bernoulli pressure, which is a quadratic function of the velocity of the liquid.

Formula (3) is interesting from several standpoints. In the first place, in the linear approximation the stability of the film does not react to the velocity u. Indeed, the stability criterion is determined by the condition that the radicand in expression (3) vanish. This gives

$$\sigma_*^4 = \kappa^2 \alpha^2, \quad q_* = 2\pi \sigma_*^2 / \alpha, \quad \kappa^2 = \rho g / \alpha, \tag{4}$$

a result known from Ref. 6 for a liquid which is stationary on average.

In the second place, the frequencies from Eq. (3) are sensitive to the sign of the product qu, so that actually the dispersion relation in dimensionless form is

$$\omega_{\pm}^j = jqu \pm [q^3 + q^2 - 2\delta^2 q^2]^{1/2}, \quad j = \pm. \tag{5}$$

Here u and q are dimensionless moduli of the velocity and wave number. Moreover,

$$q = q/\kappa, \quad \kappa^2 = \rho g / \alpha, \quad u = u(\kappa/g)^{1/2}, \quad \text{and} \quad \delta = \sigma/\sigma_*, \tag{6}$$

where σ_* is given in Eq. (4).

The frequencies $\omega_+^+(q)$ or $\omega_-^-(q)$ have a constant sign for all q. For example, the mode $\omega_+^+(q)$ given by

$$\omega_+^+(q,u,\delta) = +qu + [q^3 + q^2 - 2\delta^2 q^2]^{1/2}, \quad \omega_+^+(1,u,1) = +u \tag{7}$$

is presented in Fig. 2 (top) for $u = 0.1$ and two values of δ: $\delta = 0.96$ and 1.00. The set of these lines is labeled with a general plus sign. The analogous mode $\omega_-^-(q)$ in the lower half of the figure is labeled with a minus sign.

The combinations $\omega_+^-(q)$ and $\omega_-^+(q)$ are sign-variable. The function

$$\omega_+^-(q,\delta) = -qu + [q^3 + q^2 - 2\delta^2 q^2]^{1/2}, \quad \omega_+^-(1,u,1) = -u \tag{8}$$

for the same numbers as in Eq. (5b) is displayed in Fig. 3 with the signs $(+-)$. The $(-+)$ mode is also present here. It is obtained from Eq. (8) by making the sign change $+ \to -$.

It is obvious that the degeneracy of the sign-variable modes should be lifted at the crossing points of the spectrum. But this is now a problem of the nonlinear theory of oscillations of a charged helium surface.
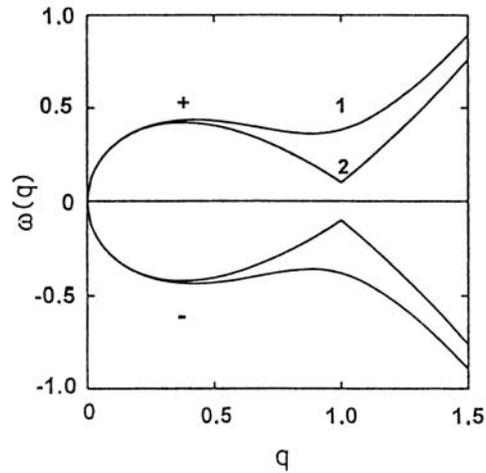
FIG. 2. The frequencies $\omega_+^+(q)$ and $\omega_-^-(q)$ (7) for $u=0.2$ and two values of δ: 1 — $\delta=0.96$, 2 — $\delta=1.00$. The first frequency is labeled with a general plus sign. The analogous modes $\omega_-^-(q)$ in the lower half of the figure labeled with a minus sign.

In summary, in the present letter a novel possibility of producing conditions for studying the relative motion between a 2D electron system and a liquid substrate is discussed. In the linear approximation this motion leads to a Doppler deformation of the dispersion relation for oscillations of a charged liquid helium surface. The most interesting nonlinear effect in prospect is the behavior of an electron crystal on a moving substrate.
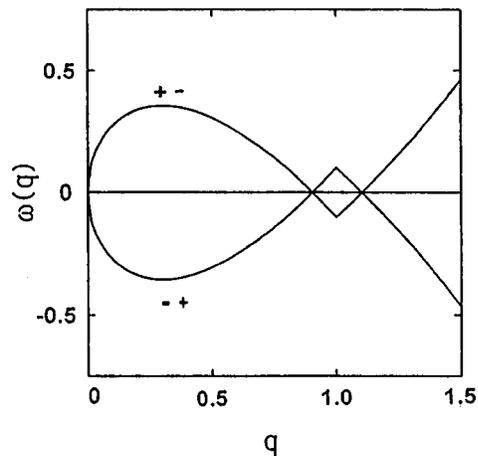
FIG. 3. The combinations $\omega_+^-(q)$ and $\omega_-^+(q)$ for the same numbers as in Fig. 2. The mode (8) is denoted by $(+-)$ signs. The mode $(-+)$, determined from Eq. (8) by making the substitution $+\rightarrow-$, is also shown here.

[a] e-mail: shikin@issp.ac.ru

[1] K. Shirahama and K. Kono, Phys. Rev. Lett. **74**, 781 (1995).
[2] K. Kono and K. Shirahama, Surf. Sci. **361-362**, 826 (1996).
[3] A. Kristensen, K. Djerfi, P. Fazooni *et al.*, Phys. Rev. Lett. **77**, 1350 (1996).
[4] V. B. Shikin and Yu. P. Monarkha, *Two-Dimensional Charged Systems in Helium* [in Russian], Nauka, Moscow, 1989.
[5] M. A. Lavrent'ev and B. V. Shabat, *Methods of the Theory of Functions of a Complex Variable* [in Russian], Fizmatgiz, Moscow, 1958.
[6] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed., rev. and enl., by E. M. Lifshitz and L. P. Pitaevskii, Pergamon Press, Oxford, 1984 [Russian original, Nauka, Moscow, 1982].

Translated by M. E. Alferieff

Effect of a magnetic field on thermally activated tunneling ionization of impurity centers in semiconductors

V. I. Perel' and I. N. Yassievich[a]

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

An expression for the probability of thermally activated tunneling ionization in an electric field in the presence of a magnetic field is obtained. It is shown that the logarithm of the ionization probability is proportional to the squared electric field, and the coefficient of proportionality decreases with increasing magnetic field. © *1998 American Institute of Physics.* [S0021-3640(98)00922-0]

PACS numbers: 79.70.+q, 61.72.−y

The effect of an electric field on the thermal ionization of impurity centers in semiconductors was previously attributed to the Poole–Frenkel effect, i.e., a decrease in the ionization energy of attracting Coulomb centers in an electric field. In recent years it has been shown theoretically and experimentally that in sufficiently strong electric fields thermally activated tunneling of charge carriers accompanied by a multiphonon transition plays the dominant role in the ionization process.[1–4] The effect of a magnetic field on this process is studied in the present letter.

Consider an impurity center on which an electron is present in a bound state. In the equilibrium configuration of the lattice, a potential well is present for the electron on the center. Lattice vibrations change this configuration, and as a result the energy of a localized electron changes. For simplicity, we shall assume that one mode of the local vibrations (ordinarily, this is the ''breathing'' mode), which is described by the configurational coordinate x, plays the dominant role in the motion of the electronic level. Ordinarily, two adiabatic potentials are introduced: $U_1(x)$, the vibrational potential in the presence of a bound electron, and $U_2(x)$, which corresponds to an ionized impurity plus a free electron with zero kinetic energy. If the free electron has energy ε, then the corresponding adiabatic potential is $U_{2\varepsilon} = U_2(x) + \varepsilon$. In an electric field the energy of a free electron can be negative (see Fig. 1).

It can be assumed that the thermal ionization process under the influence of an electric field occurs in three stages:

1) As a result of thermal excitation the system is in a vibrational level E_1 in the potential $U_1(x)$ (see Fig. 1);

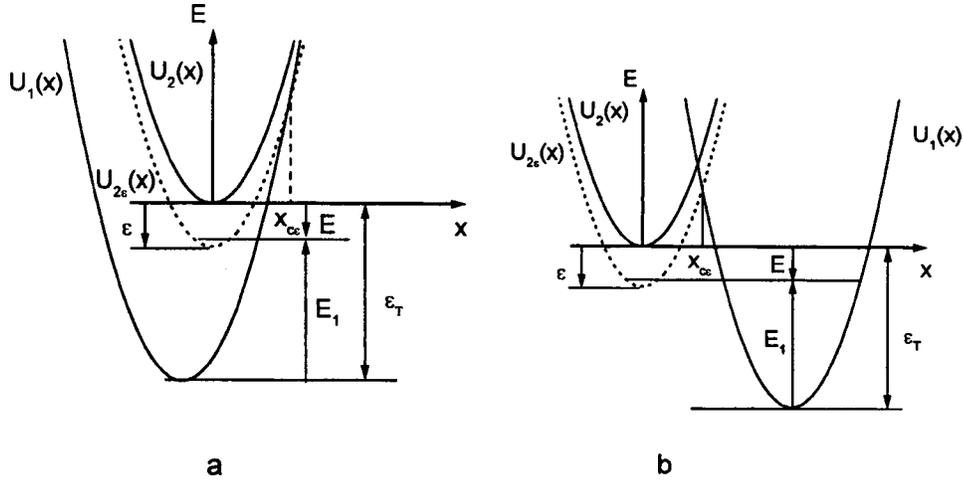2) tunneling rearrangement of the vibrational system to the adiabatic potential

FIG. 1. Diagram of the adiabatic potentials, illustrating tunneling rearrangement of the vibrational system accompanying ionization: a) and b) correspond to two possible arrangements of the potentials (autolocalization occurs in the case b). Explanations are given in the text.

$U_{2\varepsilon}(x)$, corresponding to negative electron energy ε, occurs; and,

3) an electron tunnels out of the well into a free state with energy $\varepsilon < 0$.

The first two processes occur without the participation of an electric field, while the third process occurs without the participation of vibrations.

In the quasiclassical approximation the ionization probability is proportional to the expression

$$\exp\left(-\frac{E_1}{kT}\right)\exp[-2(S_{2\varepsilon}-S_{1\varepsilon})]\exp[-2S_e(\varepsilon)], \tag{1}$$

where

$$S_{1\varepsilon}=\frac{\sqrt{2M}}{\hbar}\int_{a_1}^{x_{c\varepsilon}}\sqrt{U_1(x)-E}\,dx, \tag{2}$$

$$S_{2\varepsilon}=\frac{\sqrt{2M}}{\hbar}\int_{a_{2\varepsilon}}^{x_{c\varepsilon}}\sqrt{U_{2\varepsilon}(x)-E}\,dx. \tag{3}$$

The energy E is measured from the bottom of the potential $U_2(x)$: $E=E_1-\varepsilon_T=E_2+\varepsilon$, E_1 and E_2 are vibrational energies in the potentials $U_1(x)$ and $U_2(x)$, ε_T is the thermal ionization energy, and M is the mass associated with the chosen vibrational mode. The integration limits in Eqs. (2) and (3) are the turning points a_1 and $a_{2\epsilon}$ in the potentials $U_1(x)$ and $U_{2\varepsilon}(x)$ and the point $x_{c\varepsilon}$ where these potentials meet. The three cofactors in Eq. (1) are the probabilities of the three processes enumerated above. The last factor $\exp[-2S_e(\varepsilon)]$ determines the probability of electron tunneling with energy $\varepsilon < 0$. The total ionization probability is obtained by integrating the product (1) over the electronic energy ε and the vibrational energy E_1. Calculation of this integral by the saddle-point method yields two equations for the optimal values E_{1m} and ε_m:

$$\tau_{2\varepsilon} - \tau_{1\varepsilon} = \hbar/2kT, \quad \tau_{2\varepsilon} = \tau_e, \tag{4}$$

where

$$\tau_{1\varepsilon} = -\hbar\frac{\partial S_{1\varepsilon}}{\partial E}, \quad \tau_{2\varepsilon} = -\hbar\frac{\partial S_{2\varepsilon}}{\partial E}, \quad \tau_e = -\hbar\frac{\partial S_e(\varepsilon)}{\partial \varepsilon}, \tag{5}$$

and $E = E_{1m} - \varepsilon_T, \varepsilon = \varepsilon_m$. The times $\tau_{1\varepsilon}, \tau_{2\varepsilon}$, and τ_e are the tunneling times of the vibrational system in the potentials U_1 and $U_{2\varepsilon}$ and the electron tunneling time, respectively.

The scheme described above for calculating the ionization probability has been used previously to study ionization in dc[1] and ac[2] electric fields (see also Ref. 3). The results agree well with numerous experiments.[3,4] Here we shall employ this approach to study the effect of a magnetic field on thermal ionization in a constant electric field. We shall employ the results of Refs. 5 and 6, where the probability of electron tunneling in the presence of electric and magnetic fields was found in the quasiclassical approximation. The result obtained in these works can be written as

$$2S_e(\varepsilon) = \frac{2}{3}\frac{(2|\varepsilon|m)^{3/2}}{Fm\hbar}g(\gamma,\theta), \tag{6}$$

$$\gamma = \sqrt{2|\varepsilon|m}\,\frac{\Omega}{F}, \tag{7}$$

where m is the electron effective mass, $\Omega = eH/mc$ is the cyclotron frequency, $F = e\mathcal{E}$ is the force acting on an electron in an electric field of intensity $\mathcal{E}$. The function $g(\gamma,\theta)$, which depends on the angle θ between the directions of the electric and magnetic fields, is given by the expression

$$g(\gamma,\theta) = \frac{3}{2}\beta\left[1 - \frac{\sqrt{\beta^2-1}}{\gamma}\sin\theta - \frac{1}{3}\beta^2\cos^2\theta\right], \tag{8}$$

where $\beta > 0$ and is related with the parameter γ by the equation

$$\beta^2 - \sin^2\theta\left[\beta\coth(\beta\gamma) - \frac{1}{\gamma}\right]^2 = 1. \tag{9}$$

From Eqs. (5)–(8) we find

$$\tau_e = \frac{\gamma}{\Omega}\left[g(\gamma,\theta) + \frac{1}{3}\gamma\frac{\partial g(\gamma\theta)}{\partial\gamma}\right]. \tag{10}$$

For sufficiently weak fields, such that $|\varepsilon_m| \ll \varepsilon_T$, the difference $S_{2\varepsilon} - S_{1\varepsilon}$ in Eq. (1) can be expanded in a series in powers of ε and only the first term in the expansion need be retained. Then

$$S_{2\varepsilon} - S_{1\varepsilon} = S_2 - S_1 + \varepsilon\tau_2/\hbar, \tag{11}$$

where S_2, S_1, and τ_2 are the values of $S_{2\varepsilon}, S_{1\varepsilon}$, and $\tau_{2\varepsilon}$ at $\varepsilon = 0$. In Eqs. (4) ε can also be set to zero in the tunneling times $\tau_{1\varepsilon}$ and $\tau_{2\varepsilon}$ of the vibrational system. Then the first of Eqs. (4) determines E_{1m} and gives for the same value as in the absence of an electric field. As a result we obtain that, to exponential accuracy, the ionization probability in an electric field is given by the expression
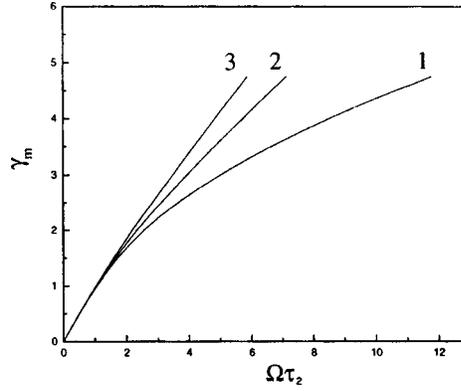
FIG. 2. γ_m versus $\Omega\tau_2$ for three values of the angle θ between the directions of the electric and magnetic fields: $1 — \theta = \pi/2, 2 — \theta = \pi/3, 3 — \theta = \pi/4$.

$$\mathbf{e}(F) = \mathbf{e}(0)\exp\left[\frac{2|\varepsilon_m|\tau_2}{\hbar} - 2S_e(|\varepsilon_m|)\right], \tag{12}$$

where the optimal energy ε_m for electron tunneling is determined by the second of Eqs. (4), where $\tau_{2\varepsilon}$ is replaced by τ_2, and expression (10) should be used for τ_e. Since the parameter γ and the energy ε are related by Eq. (7), the optimal tunneling energy ε_m corresponds to the optimal value of γ_m, for which we have finally the equation

$$\gamma_m\left[g(\gamma_m, \theta) + \frac{1}{3}\gamma_m\frac{\partial g(\gamma_m, \theta)}{\partial\gamma_m}\right] = \Omega\tau_2. \tag{13}$$

Formula (12) then becomes

$$\mathbf{e}(F) = \mathbf{e}(0)\exp(F^2/F_c^2), \tag{14}$$

where it is convenient to express F_c in terms of the auxiliary quantity τ_2^* as

$$F_c^2 = \frac{3m\hbar}{(\tau_2^*)^3} \quad \text{and} \quad (\tau_2^*)^3 = \tau_2^3\left(\frac{\gamma_m}{\Omega\tau_2}\right)^3\left(3\frac{\Omega\tau_2}{\gamma_m} - 2g(\gamma_m, \theta)\right). \tag{15}$$

As $H \to 0$ we have $g(\gamma, \theta) \to 1$ and, as is evident from Eq. (13), $\gamma_m/\Omega\tau_2 \to 1$. Therefore $\tau_2^* \to \tau_2$, and we obtain the well-known result of Refs. 1 and 3 for ionization in a constant electric field in the absence of a magnetic field. According to Eq. (4), the time τ_2 is determined by the relation

$$\tau_2 = \hbar/2kT + \tau_1, \tag{16}$$

where τ_1 is essentially independent of temperature.[3]

We shall now consider the limiting cases, confining our attention to mutually perpendicular electric and magnetic fields ($\theta = \pi/2$). In the case of a strong magnetic field $\gamma \gg 1$[6]

$$g(\gamma, \pi/2) = 3/8\gamma. \tag{17}$$
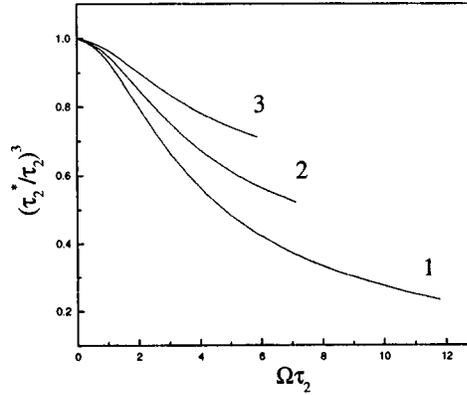
In this limit we obtain from Eq. (13)

FIG. 3. $(\tau_2^*/\tau_2)^3$ versus $\Omega\tau_2$ for three values of the angle θ between the directions of the electric and magnetic fields: $1 — \theta = \pi/2$, $2 — \theta = \pi/3$, $3 — \theta = \pi/4$.

$$\gamma_m = \sqrt{2\Omega\tau_2}. \tag{18}$$

Therefore the condition $\gamma_m \gg 1$ means that $\Omega\tau_2 \gg 1$. Under this condition we have for the parameter τ_2^* from Eq. (15)

$$\tau_2^{*3} = 3\tau_2^3/\Omega\tau_2. \tag{19}$$

The magnetic field, as expected, slows down the thermally activated tunneling ionization process.

In the weak magnetic field limit, $\Omega\tau_2 \ll 1$, it is easy to obtain from Eqs. (7), (9), (13), and (15)

$$\gamma_m = \Omega\tau_2(1 - 1/18(\Omega\tau_2)^2) \quad \text{and} \quad (\tau_2^*)^3 = \tau_2^3(1 - 1/15(\Omega\tau_2)^2). \tag{20}$$

The dependence of γ_m and $(\tau_2^*/\tau_2)^3$ on the parameter $\Omega\tau_2$ in the general case is displayed in Figs. 2 and 3 for $\theta = \pi/2, \pi/3$, and $\pi/4$.

In summary, it has been shown in this letter that a magnetic field decreases the ionization rate without changing the character of the electric field dependence of the probability of thermally activated tunneling ionization of impurity centers. This effect is determined by the parameter $\Omega\tau_2$, where Ω is the cyclotron frequency and τ_2 is determined by the tunneling rearrangement time of the vibrational system of centers in the process of ionization.

[a]e-mail: Irina.Yassievich@pop.ioffe.rssi.ru

[1] V. Karpus and V. I. Perel', JETP Lett. **42**, 497 (1985).

[2] S. D. Ganichev, E. Ziemann, Th. Gleim *et al.*, Phys. Rev. Lett. **80**, 2409 (1998).

[3] V. N. Abakumov, V. I. Perel', and I. N. Yassievich, in *Nonradiative Recombination in Semiconductors*, Vol. 33 of *Modern Problems in Condensed Matter Sciences,* edited by V. M. Agranovich and A. A. Maradudin, Amsterdam, North-Holland, 1991.

[4] S. D. Ganichev, I. N. Yassievich, and W. Prettl, Fiz. Tverd. Tela (Leningrad) **39**, 1905 (1997) [Sov. Phys. Solid State **39**, 1703 (1997)].

[5] L. P. Kotova, A. M. Perelomov, and V. S. Popov, Zh. Éksp. Teor. Fiz. **54**, 1151 (1968) [Sov. Phys. JETP **27**, 616 (1968)].

[6] V. S. Popov and A. V. Sergeev, Zh. Éksp. Teor. Fiz. **113**, 2047 (1998) [JETP **86**, 1122 (1998)].

Electron spin resonance in GaSb−InAs−GaSb semimetal quantum wells

S. A. Emel'yanov,[a] Ya. V. Terent'ev, A. P. Dmitriev, and B. Ya. Mel'tser

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021
St. Petersburg, Russia*

Resonance of the photo-emf induced by far-IR light in the presence of
a magnetic field oriented parallel to the plane of the well is observed in
a GaSb–InAs–GaSb semimetal quantum well. It is inferred that the
effect is due to optical transitions between sublevels of the first size-
quantization level. © *1998 American Institute of Physics.*
[S0021-3640(98)01022-6]

PACS numbers: 76.30.−v, 73.61.Ey

Spin resonance has thus far been observed only in structures of the type GaAs–AlGaAs or GaInAs–InP, where the two-dimensional electron gas is weakly degenerate. Resonance was observed in the form of a small microwave-induced feature in the conductivity against the background of Shubnikov−de Haas oscillations (see, for example, Refs. 1 and 2).

We report here the observation of electron spin resonance in GaSb−InAs−GaSb semimetal quantum wells. Resonance was observed in the course of an investigation of the photo-emf induced by far-IR light in the presence of an external magnetic field oriented parallel to the plane of the well.

The experiments were performed on structures with single 20-nm-thick MBE-grown quantum wells. To eliminate effects due to hybridization of the conduction band of InAs with the valence band of GaSb, the well was separated from the GaSb layers by a 10-nm-thick AlSb barrier. A characteristic density and mobility of the electrons in the well were, respectively, 4×10^{12} cm^{-2} and 7000 cm^2/V·s at $T = 300$ K and 2.5×10^{12} cm^{-2} and 9000 cm^2/V·s at $T = 77$ K. A strong built-in electric field (up to 10^5 V/cm) was present in the well. The existence of the field is characteristic for structures of the type used in Ref. 3.

Pulsed NH$_3$ and D$_2$O gas lasers optically pumped by CO$_2$ laser served as the source of radiation. The wavelength of the light, the pulse duration, and the maximum intensity of the radiation were equal to 90.6 μm, 40 ns, and 1500 W/cm^2 for the NH$_3$ laser and 385 μm, 100 ns, and 100 W/cm^2 for the D$_2$O laser. The experiments were performed at $T = 4.2$ K in the presence of an external magnetic field oriented parallel to the plane of the well. Unpolarized radiation was directed along the normal to the surface of the sample. The light-induced photocurrent arising in the plane of the well in a direction perpendicular to the magnetic field was measured. The kinetics of the measured emf corresponded to
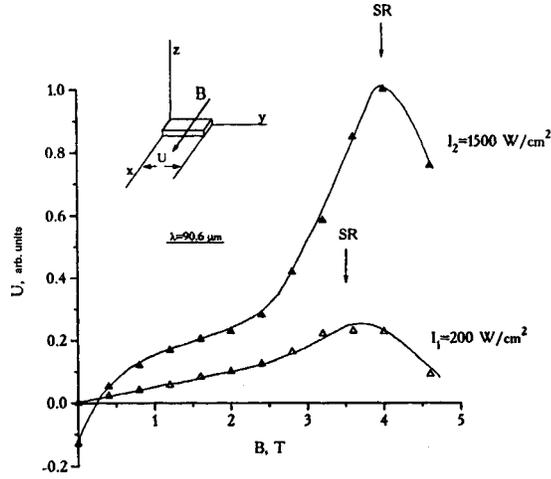
FIG. 1. Photo-emf versus the magnetic field for 90.6 μm radiation and different exciting radiation intensities. Inset: Geometry of the experiment.

that of the exciting laser pulse. The geometry of the experiment is shown in the inset in Fig. 1.

Figure 1 shows experimental curves of the photo-emf versus the magnetic field. The curves were obtained at $\lambda = 90.6\,\mu$m for two different radiation intensities. One can see that a resonance of the photo-emf, which shifts with increasing excitation intensity in the direction of high magnetic fields, is present. Reversal of the direction of the light does not change the sign of the emf. We attribute the observed resonance to optical transitions of electrons between spin sublevels of the first size-quantization level in the well. It is known that in quantum wells with a built-in electric field, even in the absence of an external magnetic field, the spin degeneracy is removed as a result of the spin-orbit interaction (so-called zero spin splitting). The energy spectrum of electrons in the plane of a two-dimensional gas (XY) has the form[4]

$$\epsilon^{\pm} = \frac{\hbar^2 k^2}{2m^*} \pm ak, \tag{1}$$

where a is the spin-orbit interaction constant, which depends on the magnitude of the built-in field. Therefore, even in the absence of an external magnetic field, the energy spectrum in the XY plane consists of two branches corresponding to two opposite directions of the spin — parallel or antiparallel to the vector of the magnetic field arising in the coordinate system of an electron moving in a strong built-in electric field. It is obvious that only electrons whose energy is close to the Fermi energy can undergo direct optical transitions. It follows from Eq. (1) that the zero spin splitting near the Fermi level is $2\alpha k_F$. To obtain a rough estimate of the energy $\hbar\omega_0$ of a photon which can be resonantly absorbed even in a zero magnetic field, we shall employ the spin-orbit splitting constant obtained from magnetotransport measurements for structures of this type:[3] $\alpha = 0.9 \times 10^{-9}$ eV·cm. Neglecting the nonparabolic nature of the spectrum and setting $\epsilon_F = 70$ meV and $m^* = 0.056m_0$ (the electron effective mass near the Fermi level[5]), we
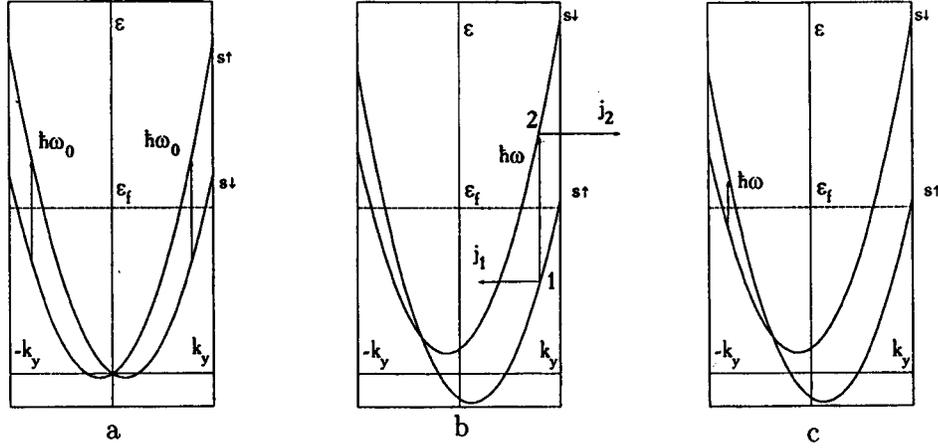
FIG. 2. a) Energy spectrum of electrons in a quantum well with a built-in electric field. The vertical arrows show the possible resonance optical transitions in a zero magnetic field. b) Energy spectrum of electrons in the direction Y in the plane of the quantum well in the presence of an external magnetic field directed along X. The vertical arrow shows the resonance optical transition accompanying the absorption of 90.6 μm light. The horizontal arrows indicate the direction of the elementary currents j_1 and j_2 arising as a result of this transition. c) Energy spectrum and scheme of resonance optical transitions accompanying the absorption of 385 μm light.

obtain $\hbar\omega_0=2\alpha k_F\sim 6$ meV. This corresponds to the energy of a far-IR photon. The energy spectrum and the scheme of possible optical transitions in a zero external magnetic field are presented in Fig. 2a.

When an external magnetic field B is applied in the plane of a two-dimensional gas, for example, along the X axis, the spectrum becomes

$$\epsilon^\pm = \frac{\hbar^2 k^2}{2m^*} \pm \sqrt{a^2 k_x^2 + \left(a k_y + \frac{1}{2}g\mu_B B\right)^2}. \tag{2}$$

Now the spins are oriented in the direction of the resultant magnetic field, which is the sum of the external magnetic field and the field arising due to the motion of an electron in the strong built-in electric field. One can see from Eq. (2) that the spectrum is now asymmetric in the Y direction (Fig. 2b).

Optical transitions are possible only for electrons with energy close to the Fermi energy. For such electrons $k_x^2+k_y^2=k_F^2$. When Eq. (2) is taken into account, the condition for a resonance at frequency ω has the form

$$(\hbar\omega)^2 = (2ak_F)^2 + (g\mu_B B)^2 + 4g\mu_B Bak_y, \tag{3}$$

where k_y can assume values from $-k_F$ to k_F. It follows from Eq. (3) that resonance absorption of light can occur in the range of frequencies of the exciting radiation from

$$\omega_{\min} = \frac{1}{\hbar}|2ak_F - g\mu_B B| \quad \text{to} \quad \omega_{\max} = \frac{1}{\hbar}|2ak_F + g\mu_B B|.$$

We shall now consider the situation occurring in an experiment where the photon energy is fixed (in our case $\hbar\omega=13.7$ meV) and B increases from zero. Resonance

absorption first occurs at $B^* = (g\mu_B)^{-1}(\hbar\omega - 2ak_F)$. In this case electrons with $k_y \approx k_F$ participate in absorption. The corresponding optical transition is shown in Fig. 2b. As B increases further, the resonance conditions will be satisfied even for electrons with $k_x \neq 0$ and, correspondingly, $k_y < k_F$. As a result, in the actual range of magnetic fields the resonance absorption contour has the form of a diffuse step. However, it is apparently very difficult to observe such a resonance experimentally because of the relatively large nonresonance absorption.

Let us now turn to a discussion of the experimentally observed resonance photocurrent. We shall consider first the case of magnetic fields close to B^*. This current is the difference between two elementary currents j_1 and j_2 which form with an optical transition $1 \rightarrow 2$ (see Fig. 2b). The currents j_1 and j_2 are not equal to one another, since the electron velocity, the electron momentum relaxation time, and the electron effective mass at the points 1 and 2 are different. For this reason, in the absence of accidental compensation, a nonzero net current will occur. It is also evident from Fig. 2b that an important feature of the scheme of the optical transitions occurring in the experimental system is that the transition of an electron with momentum $(k_x; k_y)$ is not accompanied by a symmetric transition of an electron with momentum $(k_x; -k_y)$ or close to it, as ordinarily happens when fast-response currents form (for example, the so-called entrainment currents). In our view this is the reason why in an experiment the resonance photocurrent can be distinguished against the background of a nonresonance photocurrent.

We shall now analyze the form of the resonance curve of the photo-emf. Using the relations (2) and (3), it is not difficult to show that under conditions such that $\hbar\omega$ is somewhat greater than $2ak_F$ the difference of the projections of the velocities at the points 2 and 1 on the Y axis will decrease as the magnetic field increases above B^*. In turn, this should decrease the net current. Therefore it can be expected that in contrast to resonance absorption the resonance photo-emf will be a certain curve with a maximum at the point B^*. However, if the facts that the effective mass at the point 2 is greater than at the point 1 because of the strong nonparabolicity of the spectrum, while the relaxation time of the momentum, conversely, may be smaller (the latter apparently is valid only for sufficiently large $\hbar\omega$, so that scattering by acoustic phonons predominates), are taken into account, then the net current can not only decrease appreciably, but it can even change sign.

Following the assumption that the experimentally observed photo-emf reaches its maximum value at $B = B^*$, we can estimate the spin-orbit splitting constant. Since in the experiment $B^* \cong 3.5$ T, and assuming for a rough estimate that $g = -15$ (as in bulk InAs) and $m^* = 0.056m_0$, we obtain $a = 1.5 \times 10^{-9}$ eV·cm, which is in reasonable agreement with the results of Ref. 3. The corresponding value of $2ak_F$ is approximately 10 meV, which is in complete accord with the ideas presented above.

Let us now consider the question of the shift of the position of the resonance as the excitation intensity increases. Thus far we have assumed that only electrons with energy ϵ_F (more accurately, all electrons with energy from $\epsilon_F - \hbar\omega$ to ϵ_F) can participate in resonance absorption. This valid only for sufficiently low electron temperature, so that $kT < 2ak_F$. At the same time, it is well known that far-IR radiation heats the electronic subsystem very efficiently (see, for example, Ref. 6). For the radiation intensities realized in experiments (up to 1.5 kW/cm^2), it can be expected that the characteristic scale of the dropoff of the distribution function is several tens of meV. It is obvious that for such a
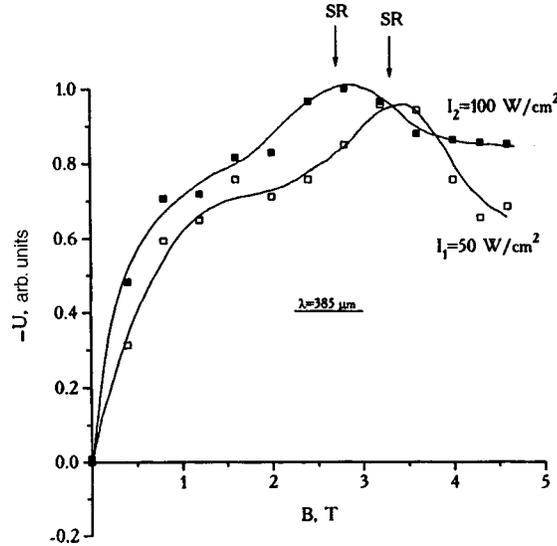
FIG. 3. Photo-emf versus the magnetic field. The curves were obtained at a wavelength of 385 μm for two different exciting radiation intensities.

diffuse distribution function, when the corresponding resonance conditions are satisfied, the strongest transitions will be optical transitions with initial energy much less than ϵ_F. This should have the effect that with increasing radiation intensity for $\hbar\omega > 2ak_F$ the position of the resonance of the photo-emf will shift in the direction of strong magnetic fields.

To check the proposed interpretation of the resonance, additional experiments were performed to measure the dependence of the photo-emf on the external magnetic field with excitation by 385 μm radiation, i.e., under conditions such that $\hbar\omega < 2ak_F$. The results are presented in Fig. 3. It is evident that the resonance photo-emf is observed here also. The following features of the resonance are of interest: (i) the position of the resonance as a function of the magnetic field (at least, for relatively low intensities) differs very little from the position of the similar resonance at $\lambda = 90.6$ μm, even though the photon energy differs by approximately a factor of 4; (ii) the resonance at $\lambda = 385$ μm is much less pronounced than at $\lambda = 90.6$ μm, and the corresponding resonance photo-emfs have opposite signs; and, (iii) as excitation intensity at $\lambda = 385$ μm increases, the resonance shifts in the direction of *weaker* magnetic fields.

In our opinion, the characteristic features enumerated above of the formation of a resonance photo-emf at $\lambda = 385$ μm agree quite well with the proposed model. Indeed, in accordance with the model the position of the resonance as a function of the magnetic field is determined essentially by the modulus of the difference of $\hbar\omega$ and $2ak_F$. Therefore resonances at frequencies ω_1 and ω_2 will be observed at close values of the magnetic fields if the value of $2ak_F$ is close to $1/2(\hbar\omega_1 + \hbar\omega_2)$. In our case this condition means that $2ak_F \approx 8.5$ meV, which agrees well with the above-presented estimate of this quantity from the position of the maximum of the resonance at $\lambda = 90.6$ μm. The fact that the amplitude of the resonance at $\lambda = 385$ μm is appreciably smaller than at $\lambda = 90.6$ μm is

likewise completely explainable, since in the first case the resonance transition is accompanied by a substantially smaller change in the electron energy and a correspondingly smaller change in the values of the electron parameters at the initial and final points of the transition. Comparing Fig. 2b and Fig. 2c, it is evident that the sign of the Y projections of the velocities at the initial and final points of the resonance transition at $\lambda = 385 \, \mu$m is opposite to that in the case of excitation at $\lambda = 90.6 \, \mu$m. Apparently, this also explains the difference in the signs of the resonance photo-emf. As far as the shift of the position of the resonance is concerned, it is not difficult to show that for $\hbar \omega < 2 a k_F$ resonance conditions for electrons with energy less than ϵ_F are realized for lower magnetic fields than in the case $\hbar \omega > 2 a k_F$. In keeping with the ideas set forth above, this means that as the excitation intensity increases, the resonance should shift in the direction of weaker fields, as is observed experimentally.

In conclusion, we shall estimate the absolute magnitude of the resonance current. Neglecting for purposes of estimation the difference in the masses and relaxation times at the points 1 and 2, it is easy to obtain for j the expression

$$j = \frac{2}{\pi} \frac{e}{\hbar^2} \int d^2 k \, \tau M^2 \delta(\epsilon_+(k) - \epsilon(k) - \hbar \omega) \theta(\epsilon_+(k) - \epsilon_F) \theta(\epsilon_F - \epsilon_-(k))$$

$$\times \frac{a\left(ak_y + \frac{1}{2} g \mu_B B\right)}{\sqrt{a^2 k_x^2 + \left(ak_y + \frac{1}{2} g \mu_B B\right)^2}}, \tag{4}$$

where M is the matrix element of the optical transition, τ is the momentum relaxation time, $\theta(\epsilon - \epsilon_F)$ is the Fermi distribution function, $\theta(\epsilon - \epsilon_F) = 1$ for $\epsilon < \epsilon_F$ and $\theta(\epsilon - \epsilon_F) = 0$ for $\epsilon > \epsilon_F$. For unpolarized radiation the squared matrix element of a resonance transition can be expressed as $M^2 = e^2 E_0^2 a^2 / 8 \hbar^2 \omega^2$, where E_0 is the amplitude of the electric field of the light wave.

Generally speaking, the calculation of the current density using Eq. (4) requires numerical integration, but in the present paper we shall only obtain an order of magnitude estimate of the maximum current. For $\lambda = 90.6 \, \mu$m this estimate gives 1 μA/cm, while experimentally the surface density of the resonance current at this wavelength with radiation intensity 200 W/cm^2 was equal to approximately 4 μA/cm. In our opinion, this likewise attests to the proposed interpretation of the experimental results.

[a]e-mail: sergey.emelyanov@pop.ioffe.rssi.ru

[1] M. Dobers, K. v. Klitzing, and G. Weimann, Solid State Commun. **70**, 41 (1989).
[2] M. Dobers, J. P. Vieren, Y. Guldner *et al.*, Phys. Rev. B **40**, 8075 (1989).
[3] J. Luo, H. Munekata, F. F. Fang, and P. J. Stiles, Phys. Rev. B **38**, 10142 (1988), Phys. Rev. B **41**, 7685 (1990).
[4] Yu. A. Bychkov and É. I. Rashba, JETP Lett. **39**, 78 (1984).

[5] L. S. Kim, H. D. Drew, H. Munekata *et al.*, Solid State Commun. **66**, 873 (1988).
[6] S. D. Ganichev, A. P. Dmitriev, S. A. Emel'yanov *et al.* Zh. Éksp. Teor. Fiz. **90**, 445 (1986) [Sov. Phys. JETP **63**, 256 (1986)].

Translated by M. E. Alferieff

Low-temperature anomalies of the Hall coefficient in FeSi

N. E. Sluchanko, V. V. Glushkov, S. V. Demishev, M. V. Kondrin,
K. M. Petukhov, A. A. Pronin, and N. A. Samarin

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia

Y. Bruynseraede and V. V. Moshchalkov

Laboratory voor Vaste-Stoffysica en Magnetisme, K. U. Leuven, B-3001 Leuven, Belgium

A. A. Menovsky

*Van der Waals–Zeeman Laboratory, University of Amsterdam, 1018 XE Amsterdam,
The Netherlands*

The temperature dependence of the Hall coefficient in the interval 1.8–
300 K is investigated in detail in high-quality single-crystal samples of
a Kondo insulator — iron monosilicide. It is established that the pa-
rameter $R_H(T, H = 12.5 \, \text{kOe})$ changes sign twice in the temperature in-
terval employed, and at temperatures below $T_m \approx 7$ K an anomalous
(magnetic) component appears in the angular and field dependences of
the Hall voltage. The results of the experimental investigations of
$R_H(T, H_0)$ in FeSi are discussed on the basis of the phase diagram in
the model of an excitonic insulator. © *1998 American Institute of
Physics.* [S0021-3640(98)01122-0]

PACS numbers: 72.20.My

1. In recent years it has been common practice to include iron monosilicide among
the compounds referred to as Kondo insulators.[1,2] However, the diversity of the theoreti-
cal approaches used in this case by different authors[1–5] and the absence of a single,
generally accepted model suitable for describing FeSi leaves open the question of the
nature of the ground state and the characteristic behavior of the physical parameters in
this material.

Among the problems involved in investigations of FeSi, it should be especially
noted that there are no reliable and accurate experimental data on the behavior of the Hall
coefficient $R_H(H, T)$ in this narrow-gap ($E_g \approx 60$ meV)[6] almost-magnetic semiconductor.
The comparatively recent[7,8] measurements of $R_H(H, T)$ pertain to substantially different
temperature intervals ($T \lesssim 55$ K in Ref. 7 and $T \gtrsim 20$ K in Ref. 8), and they do not agree
even with respect to the sign of the Hall coefficient. At the same time accurate data on
$R_H(T, H)$ in relation to measurements of the thermopower (Seebeck coefficient) and the
resistance of a narrow-gap semiconductor will make it possible to determine the structure
and character of the ''impurity'' states in the gap and the mechanisms of scattering of
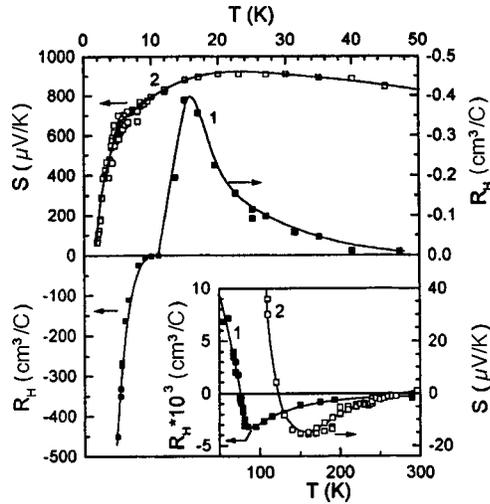charge carriers in FeSi.

FIG. 1. Temperature dependences of the Hall coefficient R_H (*1*) and thermopower S (*2*) of a FeSi single crystal.

In this connection, our aim in the present investigation is to obtain detailed information about the characteristics of charge carriers in the region of intrinsic and extrinsic conductivity in FeSi on the basis of measurements of the Hall coefficients, the thermopower, and the resistance in the temperature interval 1.8–300 K. Since in Ref. 8 the appearance of hysteresis in the field-dependences of the Hall voltage in FeSi was observed at ultralow temperatures, to determine the character of the ground state it is of special interest in the present investigation to study also the characteristic behavior of $R_H(T,H)$ at liquid-helium temperatures. The procedure used of the thermoelectric measurements is similar to that discussed in Ref. 9. To investigate $R_H(H_0,T)$, in addition to the standard technique,[10] we employed an automated module for controlling the rotation of the sample in a magnetic field and a high-precision temperature controller with a novel construction. The measurements were performed on high-quality single-crystal samples ($\rho(1.8 \text{ K})/\rho(300 \text{ K}) \geqslant 10^5$).

2. The experimental data obtained from investigations of the Hall coefficient $R_H(T)$ in the magnetic field $H = 12.5$ kOe (curve *1*) and the thermopower $S(T)$ (curve *2*) in the temperature interval 1.8–300 K are displayed in Fig. 1. According to this figure, for FeSi we observed two inversion of the sign of the Hall coefficient as the temperature was lowered from room temperature — at $T_{\text{inv1}}^H \approx 75$ K and $T_{\text{inv2}}^H \approx 12.5$ K. The existence of two linear activation sections in the function $R_H(T)$, plotted in the coordinates $\ln R_H = f(1/T)$, in the temperature intervals 80–300 K (I) and 15–60 K (II) (Fig. 2) makes it possible to determine the parameters $T_{a1}^H \approx 345$ K and $T_{a2}^H \approx 68$ K. At temperatures $T < T_{\text{inv2}}^H \approx 12.5$ K, rapid growth of the absolute value of the Hall coefficient is observed (Fig. 1); in addition, as temperature decreases, an anomalous contribution to $R_H(T,H)$, accompanied by hysteresis in the field and angular dependences of the Hall voltage, appears in the region $T \leqslant T_m \approx 7$ K (region III).
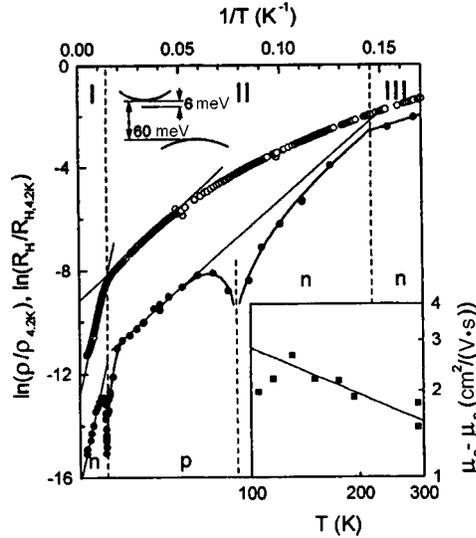
FIG. 2. Activation asymptotics of the temperature dependences of the resistivity ρ ($\bigcirc$) and Hall coefficient $R_H(H = 12.5$ kOe) ($\bullet$) of iron monosilicide. The sign of R_H is determined by the indicated type of charge carriers (p or n). Inset: Temperature dependence of the difference of the electron and hole mobilities in the region of intrinsic conductivity of FeSi (the approximation by a linear function corresponds to the law $T^{-1/2}$).

The behavior of the thermopower $S(T)$ in FeSi (curve *2* in Fig. 1) is similar to that measured in Refs. 7 and 11, and on the whole it is similar to the temperature dependence $R_H(T)$. In the temperature interval 150–300 K, replotting the experimental data in terms of the coordinates $S = f(1/T)$ yields $T^S_{a1} \approx 70$ K. We emphasize that the low-temperature maximum in $S(T)$ with a positive amplitude greater than 900 μV/K (see Fig. 1, curve *2*), following the arguments of Ref. 7, must be attributed to phonon drag in FeSi. Apparently, the appearance of an additional contribution in $S(T)$ in this temperature interval not only strongly shifts the inversion point $T^S_{\text{inv}1} \approx 123$ K $> T^H_{\text{inv}1}$ but it also makes it difficult to determine directly the activation parameters T^S_{a2} and E^S_{a2} from the function $S(T)$. Following Ref. 7, we also note the existence of a sign change in $S(T)$ near room temperature. Another feature of the behavior of $S(T)$ in FeSi observed in the present investigation is the appearance of a kink in the temperature dependence of the thermopower near $T_m \approx 7$ K (see Fig. 3c).

To obtain the additional information required to analyze the temperature dependences of the Hall coefficients and thermopower (Fig. 1, curves *1* and *2*), measurements of the dc resistance (curve *2* in Fig. 2) and the microwave conductivity (see Fig. 3b) in the temperature 1.8–300 K were also performed in the present work. The temperature dependence $\rho(T)$ agrees on the whole with the results of Refs. 7–8 and 11. Just as in the case $R_H(T)$, the curve $\rho(T)$ contains two activation sections of variation of the resistance, which are virtually identical to the intervals of activation behavior of the Hall coefficient (Fig. 2). The values of the parameters T^ρ_{a1} and T^ρ_{a2} are 336 K and 71 K, respectively, in agreement, to within the limits of the experimental error, with the values
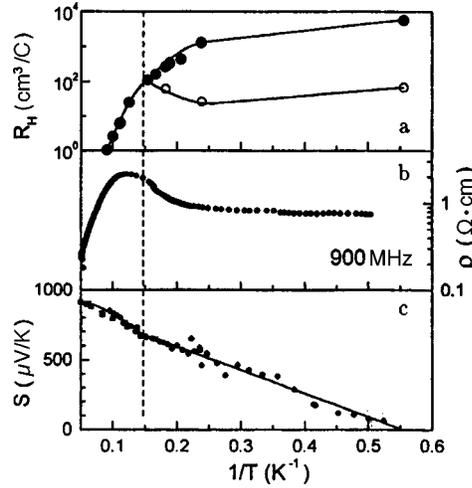
FIG. 3. Low-temperature dependences a) of the anomalous (●) and normal (□) Hall coefficients R_H, b) dynamic resistivity ρ at 900 MHz, and c) thermopower S of iron monosilicide near T_m.

of T_{a1} and T_{a2} determined from the Hall data. In contrast to the dc resistance, which is a monotonic function of temperature, we observed in the temperature dependence of the microwave resistance a feature in the form of a peak near $T_m \approx 7$ K (Fig. 3b), corresponding to the temperature at which the anomalous (magnetic) component of the Hall voltage and a kink in $S(T)$ appear.

3.A. REGION OF INTRINSIC CONDUCTIVITY $T \gtrsim 70$ K (I)

Apparently, the matched change in the parameters $\rho(T)$ and $R_H(T)$ in this temperature interval (Fig. 2) should be attributed to the activational decrease of the density of intrinsic charge carriers with temperature. The gap width estimated from the data in Fig. 2 is $E_g = 2T_{a1} \approx 60$ meV and agrees with the results of Refs. 6–8 as well as with the value of E_g obtained from calculations of the electronic structure of FeSi.[12–14] Using the expression for the thermopower of a semiconductor in the region of intrinsic conductivity

$$S = \frac{k_B}{e} \left\{ \frac{b-1}{b+1} \frac{E_g}{2k_B T} + \frac{3}{4} \ln \frac{m_n}{m_p} \right\},$$

where $b = \mu_n / \mu_p$ and μ_n, m_n, μ_p, and m_p are, respectively, the mobility and effective masses of electrons and holes, k_B is Boltzmann's constant, and e is the electron charge, we obtain the estimate $b = \mu_n / \mu_p = 1.51$. Next, to assess the character of the temperature variation of the difference of the carrier mobilities in the region of intrinsic conductivity, we shall employ the very simple relation $\mu_n - \mu_p = R_H / \rho = AT^{-\alpha}$ and the experimental data in Fig. 2. The exponent found in this manner (see inset in Fig. 2) $\alpha \approx 0.5$, together with the quite low values of $\mu_n, \mu_p \approx 4-6$ cm^2/V·s, can be attributed in our view to scattering of carriers by local spin-density fluctuations,[3,4] which increases the modulus of the thermally induced magnetic moments in FeSi in the temperature range $T \gtrsim 100$ K.

An alternative explanation of such a low charge carrier mobility can be obtained in the model[5] of charge fluctuations $3d^6Fe^{2+} \rightarrow 3d^7Fe^{1+}$ in FeSi.

B. REGION OF EXTRINSIC CONDUCTIVITY 15–60 K (II)

For such low and quite close values of the mobilities in the electronic and hole subsystems, the appearance of an additional scattering channel predominantly for charge carriers of the same type (electrons) should result in a change in the relation $b = \mu_n/\mu_p$ and a change in the sign of the parameters $R_H(T)$ and $S(T)$. The appearance of an additional level in the gap with $E_{ex} \approx 6$ meV (see also inset in Fig. 2), in our opinion, can be attributed to the appearance of a feature, corresponding to excitonic states, in the spectrum of quasiparticle excitations near the conduction-band bottom. A possible analogy with another Kondo insulator SmB_6, in which, it seems to us,[15] rapid valence fluctuations can be described quite well in the Kikoin–Mishchenko exciton–polaron model,[16] should be underscored. We also note that for both SmB_6 and FeSi the presence of congruent electron–hole sections of the Fermi surface (which appear, for example, as a result of p–d (d–f) hybridization) is a favorable factor for describing these narrow-gap semiconductors on the basis of the model of an excitonic insulator (EI).[17]

Evidently, such low electron and hole mobilities at intermediate temperatures could attest both to rapid relaxation of charge carriers as a result of scattering by spin/charge fluctuations in FeSi and to the formation of heavy fermions as a result of efficient hybridization of the p–d states in iron monosilicide. To check all this we present, assuming that excitonic states appear in the gap, a very simple estimate of the effective mass of the carriers. Using the relation $E_{ex} \approx m^* e^4/4\epsilon^2\hbar^2$ and our value $E_{ex} = 6$ meV and $\epsilon_0 = \epsilon'(\omega \rightarrow 0) \approx 200$ (Ref. 18) we have for m^* in FeSi $m^* \approx 34.5m_0$. Using the values of m^* and E_{ex}, we estimate the localization radius of ''impurity'' states as $a^* = \hbar/\sqrt{2m^*E_{ex}} \approx 4$ Å . The value obtained for the parameter a^* correlates with the result for SmB_6[15] and, in our view, is an additional argument in favor of the proposed interpretation.

C. TEMPERATURE RANGE $T \leqslant 7$ K (III)

When excitonic insulator models are used to describe the ground state of FeSi, the presence of an electronic phase transition near $T_m \approx 7$ K (see Fig. 3) could be due to, for example, a transition, on the basis of the phase diagram of an EI[19] into a state with a spin-density wave. Our measurements of the angular and field dependences of the Hall coefficient $R_H(T \leqslant 7$ K) provide additional arguments in favor of such a conclusion, but the results of these investigations fall outside the scope of the present work.

In conclusion, we note that to resolve conclusively the question of the nature of the ground in FeSi it is necessary to perform low-temperature measurements of the quasioptic and magnetooptic properties of iron monosilicide in the millimeter and submillimeter ranges.

[1] G. Aeppli and Z. Fisk, Comments Condens. Matter Phys. **16**, 155 (1992).

[2] Z. Fisk, J. L. Sarrao, J. D. Thompson *et al.*, Physica B **206−207**, 798 (1995).

[3] T. Moriya, J. Magn. Magn. Mater. **14**, 1 (1979).

[4] P. V. Gel'd, A. G. Volkov, S. V. Kortov *et al.*, Dokl. Akad. Nauk SSSR **320**, 1097 (1991) [Sov. Phys. Dokl. **36**, 698 (1991)].

[5] C. M. Varma, Phys. Rev. B **50**, 9952 (1994).

[6] D. Mandrus, J. L. Sarrao, A. Magliori *et al.*, Phys. Rev. B **51**, 4763 (1995).

[7] B. C. Sales, E. C. Jones, B. C. Chakoumakos *et al.*, Phys. Rev. B **50**, 8207 (1994).

[8] S. Paschen, E. Felder, M. A. Chernikov *et al.*, Phys. Rev. B **56**, 12916 (1997).

[9] N. E. Sluchanko, V. V. Glushkov, S. V. Demishev *et al.*, Phys. Rev. B **56**, 10816 (1997).

[10] N. E. Sluchanko, V. V. Glushkov, S. V. Demishev *et al.*, Phys. Rev. B **51**, 1112 (1995).

[11] B. Buschinger, C. Geibel, F. Steglich *et al.*, Physica B **230−232**, 784 (1997).

[12] L. I. Vinokurova, A. V. Vlasov, and É. T. Kulatov, Trudy IOFAN **32**, 26 (1991).

[13] L. Matteiss and D. Hamann, Phys. Rev. B **47**, 13114 (1993).

[14] T. Jarlborg, Phys. Rev. B **51**, 11106 (1995).

[15] N. E. Sluchanko, V. V. Glushkov, S. V. Demishev *et al.*, Zh. Éksp. Teor. Fiz. **115** (1999) [JETP], at press.

[16] K. A. Kikoin and A. S. Mishchenko, J. Phys.: Condens. Matter **7**, 307 (1995).

[17] L. V. Keldysh and Yu. V. Kopaev, Fiz. Tverd. Tela (Leningrad) **6**, 2791 (1964) [Sov. Phys. Solid State **6**, 2219 (1964)].

[18] A. Damascelli, K. Schulte, D. van der Marel *et al.*, Phys. Rev. B **55**, R4863 (1997).

[19] N. I. Kulikov and V. V. Tugushev, Usp. Fiz. Nauk **144**, 643 (1984) [Sov. Phys. Usp. **27**, 954 (1984)].

Translated by M. E. Alferieff