Weyl Anomaly in Two-Dimensional Supergravity

D. R. Karakhanyan

Yerevan Institute of Physics, Academy of Sciences of Armenia, Yerevan, 375036 Armenia

e-mail: karakhan@lx2.yerphi.am

Two-dimensional supergravity theory is analyzed in terms of Weyl group cohomologies. A covariant nonlocal expression for the effective action is constructed. As in the case of ordinary gravity, its basic nonlocal part is determined by the well-known Polyakov action $R \frac{1}{\Box} R$. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 04.65.+e; 11.25.Hf

1. In the last two decades, two-dimensional field-theory models were investigated extensively and comprehensively. Besides direct application to certain fields of physics, these models provide excellent (and often unique) opportunity for examining realistic, but considerably more complex four-dimensional models. Indeed, the phenomena such as asymptotic freedom, dimensional transmutation, etc., also arise in two-dimensional integrable theories and can be exactly described. By examining these models, we hope to find a way to construct the corresponding four-dimensional mechanism. Finally, the study of two-dimensional models may be helpful in gaining some information on the structure of four-dimensional theories. From this viewpoint, two-dimensional gravity is the most interesting two-dimensional theory.

It is well known that the Einstein–Hilbert action in two dimensions is a topological invariant proportional to the Euler characteristic of a two-dimensional manifold. Because the standard gravity Lagrangian is locally trivial (proportional to the total derivative within an arbitrary coordinate vicinity on the manifold), it is pertinent to remember Sakharov's hypothesis that the classical gravitational Lagrangian should be identical to zero while the true Hamiltonian is induced by quantum fluctuations of gravitating matter fields. This is precisely the theory to which the problem of uncritical strings was reduced after the famous work [1] by Polyakov, who subsequently connected [2] the formulation of this problem in the light-like gauge with the Wess–Zumino–Novikov–Witten model for the $SL(2, R)$ group and calculated, on this basis, some important characteristics of the string theory. David, Distler, and Kawai [3] reproduced this result directly in the conformal gauge. Later on, Knizhnik, Polyakov, and Zamolodchikov [4] proposed the method of evaluating the exponents for the dimensional transmutation induced in primary fields by their gravitational interaction in the quantum two-dimensional conformal field theory. They found that the condition for the cancellation of quantum anomalies requires the fractional dimensionality of space–time. The supersymmetric generalization of [2] was given by Polyakov and Zamolodchikov in [5].

David, Distler, and Kawai [3] have proved, in fact, that the observable parameters of the theory are independent of a regularization scheme because this scheme is determined by the gauge chosen in calculating the effective action. The equivalence of different regularization schemes was discussed in [6]. In [7], induced two-dimensional supergravity was formulated in terms of the Weyl-invariant regularization scheme.

Holomorphic properties of the effective action of induced gravity were analyzed in [8, 9]. An attempt to take into account the effect of an external curvature on the action of induced gravity was made in [10].

Fujiwara *et al.* [11] examined the relation between the super-Weyl and super-Virasoro anomalies and constructed the super-Weyl-invariant functional by canceling the super-Virasoro anomaly with the use of a nonlocal functional.

A nonlocal action depending only on the gravitational variables was presented for the induced two-dimensional supergravity by Grisaru and Xu [12]. However, the theory has not yet been formulated in a covariant form similar to the expression $R \frac{1}{\Box} R$ for the ordinary two-dimensional gravity, i.e., without involving additional variables. This work is aimed at filling this gap.

The result obtained in this work can be generalized to higher dimensions. The Weyl anomaly was analyzed for the 4D and 6D theories of ordinary gravity in [13], where Polyakov's action was generalized to these cases. However, in contrast to 2D, the number of degrees of freedom for higher dimensions exceeds the number of classical action symmetries and the total effective action cannot be reproduced by integrating the anomalous Ward identities; the dependence on the other components of the metric tensor will be expressed in terms of integration constants. For this reason, the nonlocal expression derived in such a way will be only a part corresponding to the contribution of the Weyl anomaly to the total effective action.

2. In gravity theory, graviton excitations are described by traceless metric disturbances, and the general covariance requires that the classical action be unchanged under these variations of the metric tensor. It is well-known that the general covariance in two-dimensional theory is always accompanied by the Weyl invariance: any expression written as an integral over a two-dimensional manifold of the scalar density constructed of local matter fields and metric tensor is invariant about the Weyl transformation; i.e., $g_{\alpha\beta}(x) \longrightarrow \rho(x)g_{\alpha\beta}$. Thus, the Weyl invariance implies that the action is also independent of the residual variations of the metric $g^{\alpha\beta}\delta g_{\alpha\beta}$. In other words, the general covariant action in two-dimensional theory is independent of the metric tensor, and two-dimensional space–time is locally flat or, strictly speaking, may be flattened in every coordinate vicinity through the coordinate transformation.

However, quantum fluctuations break this symmetry so that the variation of the effective action under the conformal transformations of the metric is no longer zero:

$$g^{\alpha\beta}\frac{\delta W}{\delta g_{\alpha\beta}} = R, \qquad (1)$$

where R is the curvature of the space–time manifold. It is easy to see that, being a conformal anomaly, the curvature R satisfies the self-consistency condition; i.e., under the Weyl transformations

$$\delta g_{\alpha\beta} = g_{\alpha\beta}\delta\sigma \qquad (2)$$

the curvature is transformed as

$$\delta[\sqrt{g}R] = \sqrt{g}\square\delta\sigma, \qquad (3)$$

where $\square$ is the Laplacian acting on scalar fields. Therefore, the second derivative of the effective action W

with respect to σ is symmetric:

$$\frac{\delta R(x)}{\delta\sigma(y)} = \frac{\delta R(y)}{\delta\sigma(x)}, \qquad (4)$$

i.e., R determines the self-consistent expression for the Weyl anomaly.

A graviton partner in the two-dimensional supergravity (gravitino) is described by the Rarita–Schwinger field components corresponding to spin 3/2, so that local supersymmetry implies that the action is independent of these components. On the classical level, the two-dimensional superspace is also locally flat, because the local supersymmetry in two dimensions is always accompanied by the super-Weyl symmetry, signifying that the action is independent on the components of spin 1/2 as well:

$$\delta\chi_{\alpha} = \gamma_{\alpha}\delta\lambda. \qquad (5)$$

This symmetry also breaks by the quantum corrections: the convolution of a supercurrent J^{α} theory, i.e., the variational derivative of the effective action with respect to the super-Weyl parameter, with the γ matrices

$$\delta W/\delta\lambda = \gamma_{,\alpha}J^{\alpha} \qquad (6)$$

is no longer zero, in contrast to the derivative of the classical action.

To obtain an explicit expression for the super-Weyl anomaly, one can use the condition for integrability of the effective action (Wess–Zumino self-consistency condition): the variational derivative of the curvature (Weyl anomaly) with respect to λ should be equal to the derivative of the super-Weyl anomaly with respect to σ. The λ dependence of the curvature originates from the nonminimal term in the expression for spinor connectivity:

$$\omega_{\alpha} = -e_{\alpha}^{a}\frac{\epsilon^{\mu\nu}}{e}\partial_{\mu}e_{\nu}^{a} - 2i\bar{\chi}_{\alpha}\gamma_3\gamma^{\mu}\chi_{\mu}, \qquad (7)$$

where e is the determinant of the set of coordinates. By solving this condition, one obtains the following set of anomalous Ward identities for the effective action calculated for the two-dimensional supergravity in the supercoordinate-invariant regularization:

$$g^{\alpha\beta}T_{\alpha\beta} = R,$$
$$\gamma^{\alpha}J_{\alpha} = -4i\frac{\epsilon^{\alpha\beta}}{e}\gamma_3 D_{\alpha}\chi_{\beta}, \qquad (8)$$

where $D_{\alpha}\lambda = \partial_{\alpha}\lambda + (\gamma_3\omega_{\alpha}\lambda)/2$.

Finally, the third and last component of the Wess–Zumino condition is also satisfied: the derivative of the super-Weyl anomaly (8) with respect to λ is symmetric about the change of arguments, as also is the derivative of the curvature with respect to σ. This expression for

the anomaly coincides with the result obtained by Fuji-wara *et al.* in [14]. Taking finite Weyl and super-Weyl variations of the right-hand sides of Eqs. (8), multiplying them, respectively, by $\delta\sigma$ and $\delta\lambda$, and adding the results, one finds that the total variation of the effective action is equal to the sum of the Neveu–Schwarz action and integral of $R\sigma$ and $-4i\dfrac{\epsilon^{\alpha\beta}}{e}\bar{\lambda}\gamma_3 D_\alpha\chi_\beta$. Indeed, because Eqs. (8) are self-consistent, the above-described variational equation is integrable to give

$$S(\sigma, \lambda; e_\alpha^a, \chi_\alpha) = \int e d^2 x[R\sigma - 4i\epsilon^{\alpha\beta}/e\bar{\lambda}\gamma_3 D_\alpha\chi_\beta$$
$$- 1/2 g^{\alpha\beta}\partial_\alpha\sigma\partial_\beta\sigma - i/2\bar{\lambda}\gamma^\alpha\partial_\alpha\lambda \qquad (9)$$
$$+ i\bar{\lambda}\gamma^\beta\gamma^\alpha\chi_\beta(\partial_\alpha\sigma - i/2\bar{\lambda}\chi_\alpha)].$$

This result was also obtained by Kamimura *et al.* in [15]. The functional $S(\sigma, \lambda; e_\alpha^a, \chi_\alpha)$ is cocyclic:

$$S(\sigma_1 + \sigma_2, \lambda_1 + \lambda_2; e_\alpha^a, \chi_\alpha)$$

$$= S\left(\sigma_1, e^{-\sigma_2/4}\lambda_1; e^{\sigma_2/2}e_\alpha^a, e^{\sigma_2/4}\left(\chi_\alpha + \frac{1}{4}\gamma_\alpha\lambda_2\right)\right) \quad (10)$$

$$+ S(\sigma_2, \lambda_2; e_\alpha^a, \chi_\alpha).$$

From the viewpoint of the cohomology theory, the cochains for the Weyl transformations in the theory of two-dimensional supergravity should be dimensionless local actions constructed of the gravitational (e_α^a, χ_α) and group (σ, λ) variables and having the corresponding cocyclicity properties. The variables σ and e_α^a are dimensionless, χ_α and λ have a dimensionality of 1/2, and the derivative has a dimensionality of 1.

The quantum anomaly corresponds, in quantization, to the transition from the exact representations in the space of functions defined on the phase space of classical theory to the projective representations of the state wave function in quantum theory. This anomaly is determined by the phase multiplier of the wave function (partition function) in the theory and is a 1-cocycle. Then, group multiplication immediately leads to Eq. (10) for this phase multiplier. A 0-cochain is determined by an arbitrary counterterm (it does not contain group parameters), and its addition to the effective action corresponds to the finite renormalization of the latter. An arbitrary coboundary, i.e., finite (super)Weyl variation of a 0-cochain, is a 1-cocycle. If the anomalous phase multiplier is determined by the coboundary of some local functional, the latter can be absorbed by redefining the effective action. Such an anomaly is removable and the corresponding theory is nonanomalous.

According to this definition, functional (9) is a 1-cochain. It is obvious that each 1-cochain is no more than quadratic in group variables, and the terms of the maximum power of group variables are super-Weyl-invariant.

In the case under consideration, the anomaly is unremovable and cocycle (9) can be represented in the form of a coboundary only in a broader class of nonlocal functionals:

$$S(\sigma, \lambda; e_\alpha^a, \chi_\alpha)$$
$$= W\left[e^{\sigma/2}e_\alpha^a, e^{\sigma/4}\left(\chi_\alpha + \frac{1}{4}\gamma_\alpha\lambda\right)\right] - W[e_\alpha^a, \chi_\alpha]. \qquad (11)$$

This functional is supercovariant, but is not invariant about the Weyl transformations and corresponds to the effective action calculated in the regularization scheme invariant with respect to the superdiffeomorphisms. Indeed, the desired effective action is determined by anomalous Eqs. (8). Equation (11) obviously satisfies these restrictions. The cocyclicity of S is a trivial consequence of this expression. For this reason, I will try to find the effective action on the basis of Eq. (9). However, an attempt at performing Gaussian integration over the group variables, considering its quadratic dependence on the latter, does not give the desired result, because of the presence of a mixed term $i\bar{\lambda}\gamma^\beta\gamma^\alpha\chi_\beta\partial_\alpha\sigma$ in Eq. (9). When treating Eq. (11) as an initial relation for determining $W[e_\alpha^a, \chi_\alpha]$ and trying to satisfy it by a nonlocal functional of the form *anomaly (Weyl-invariant differential operator)$^{-1}$ anomaly* [13] [with allowance made for Eq. (9)], one must take into account the arbitrariness in the choice of local counterterms. The kinetic operators appearing in the Neveu–Schwarz action as matrix elements of the matrix of second derivatives of effective action can be taken as a Weyl-invariant operator in this symbolic expression. Unfortunately, this matrix is nondiagonal. It can be diagonalized by finite renormalization (corresponding to the addition of a counterterm) of the measure of functional integral in the definition of effective action:

$$W'[e_\alpha^a, \chi_\alpha] = W[e_\alpha^a, \chi_\alpha] + 8i\int e d^2 x\bar{\chi}_\beta\gamma^\beta D^\alpha\chi_\alpha. \quad (12)$$

This renormalization clearly leads to the redefinition of currents, and the set of anomalous equations takes the form

$$\frac{\delta W'}{\delta\sigma(x)} \equiv g^{\alpha\beta}T'_{\alpha\beta} = R - 4i\nabla_\alpha(\bar{\chi}_\beta\gamma^\beta\chi^\alpha),$$
$$\frac{\delta W'}{\delta\lambda(x)} \equiv \gamma^\alpha J'_\alpha = -4iD_\alpha(\gamma^\alpha\gamma^\beta\chi_\beta). \qquad (13)$$

The cocycle arising upon the integration of this set of anomalous Ward identities is the sum of cocycle (9) and

the coboundary of the added counterterm and is linear in the group variables. Taking this fact into account, one can easily verify that the expression

$$\int e\left[(R-i/4\nabla_\alpha(\bar\chi_\mu\gamma^\mu\chi_\alpha))\frac{1}{e\square}e(R-i/4\nabla_\beta(\bar\chi_\nu\gamma^\nu\chi_\beta))\right.$$

$$+e^{3/4}(-4iD_\alpha(\bar\chi_\mu\gamma^\mu\gamma^\alpha)) \tag{14}$$

$$\left.\times\frac{1}{e^{1/2}(-1\gamma^\nu D_\nu+\chi_\nu\bar\chi_\rho\gamma^\nu\gamma^\rho)}e^{3/4}(-4iD_\beta(\gamma^\beta\gamma^\tau\chi_\tau))\right]$$

has the same transformation properties as does $W'[e,\chi]$. Indeed, when taking the Weyl coboundary (finite variation), the nonlocality cancels out through the same mechanism as in the pure boson theory for Polyakov's action. As a result, the coboundary of the above expression yields two independent cocycles:

$$S(\sigma;e_\alpha^a,\chi_\alpha)+S(\lambda;e_\alpha^a,\chi_\alpha)$$

$$=W'\left[e^{\sigma/2}e_\alpha^a,e^{\sigma/4}\left(\chi_\alpha+\frac14\gamma_\alpha\lambda\right)\right]-W'[e_\alpha^a,\chi_\alpha], \tag{15}$$

where

$$S(\sigma;e_\alpha^a,\chi_\alpha)=\int ed^2x$$

$$\times(R\sigma-1/2g^{\alpha\beta}\partial_\alpha\sigma\partial_\beta\sigma+4i\bar\chi_\beta\gamma^\beta\chi^\alpha\partial_\alpha\sigma) \tag{16}$$

is a 1-cocycle with respect to the Weyl transformations

$$S(\sigma_1+\sigma_2;e_\alpha^a,\chi_\alpha)$$

$$=S(\sigma_1;e^{\sigma_2/2}e_\alpha^a,e^{\sigma_2/4}\chi_\alpha)+S(\sigma_2;e_\alpha^a,\chi_\alpha) \tag{17}$$

and is super-Weyl-invariant. The second cocycle

$$S(\lambda;e_\alpha^a,\chi_\alpha)=\int ed^2x(i/2\bar\lambda\gamma^\alpha\partial_\alpha\lambda+2i\bar\chi_\beta\gamma^\beta D_\alpha\chi^\alpha$$

$$+2\bar\chi_\mu\gamma^\mu\gamma_3\chi_\alpha\bar\lambda\gamma_3\gamma^\beta\gamma^\alpha\chi_\beta-1/4\bar\lambda\lambda(\bar\chi_\beta\gamma^\alpha\gamma^\beta\chi^\alpha)) \tag{18}$$

is Weyl-invariant and is a 1-cocycle with respect to the super-Weyl transformations:

$$S(\lambda_1+\lambda_2;e_\alpha^a,\chi_\alpha)$$

$$=S\left(\lambda_1;e_\alpha^a,\left(\chi_\alpha+\frac14\gamma_\alpha\lambda_2\right)\right)+S(\lambda_2;e_\alpha^a,\chi_\alpha). \tag{19}$$

A closer inspection shows, however, that the last cocycle is trivial, i.e., is a coboundary of the local counterterm.

Thus, the final expression for $W[e,\chi]$ takes the form

$$W[e,\chi]=4i\int d^2x\bar\chi_\beta\gamma^\alpha\gamma^\beta D_\alpha(\gamma_\mu\chi^\mu)$$

$$+\int e(R-i/4\nabla_\alpha(\bar\chi_\mu\gamma^\mu\chi_\alpha))\frac{1}{e\square}e(R-i/4\nabla_\alpha(\bar\chi_\mu\gamma^\mu\chi_\alpha)). \tag{20}$$

This result is not surprising, because the torsion terms are known to split off in two-dimensional supergravity, and the curvature can be determined by the metric tensor alone. It is this curvature which appears in the resulting expression for the effective action.

In this interpretation, the first term in Eq. (20) reflects the contribution of the separated Fermi fields to the finite renormalization of the action.

3. Thus, dealing with the Weyl anomaly, I arrive at the following basic conclusion. The most general form of the conformal anomaly is expressed in terms of three types of solutions of the corresponding Wess–Zumino self-consistency condition: the density of the Euler characteristic of the space–time manifold, the Weyl-invariant Lagrangian densities, and the variational derivatives of the dimensionless local actions with respect to the Weyl parameter (improving terms).

Among the solutions of the self-consistency conditions, there is always a solution that has a linear (and diagonal in the supersymmetric case) finite Weyl variation. The corresponding cocycle is quadratic in the Weyl parameter, and the corresponding anomalous phase multiplier, being integrated with respect to the group variables, makes a conformal anomaly contribution to the effective action.

The above argument applies to four-dimensional supergravity as well. However, in contrast to two-dimensional supergravity, one cannot reconstruct the whole effective action in this case, because the equations for the anomaly contain information only about the anomalous dependence of the effective action on the Weyl supergroup parameters: the determinant of the metric and the Rarita–Schwinger field components corresponding to spin 1/2. After allowance made for the general covariance and local supersymmetry, the effective action still depends on five degrees of freedom of the metric and eight degrees of freedom of the Rarita–Schwinger field, which is not fixed by the anomalous equations. The corresponding terms in the effective action arise as integration constants in solving the anomalous equations.

REFERENCES

1. A. M. Polyakov, Phys. Lett. B **103B**, 207 (1981); **103B**, 211 (1981).

2. A. M. Polyakov, Mod. Phys. Lett. A **2**, 893 (1987).

3. F. David, Mod. Phys. Lett. A **3**, 1651 (1988); F. Distler and H. Kawai, Nucl. Phys. B **321**, 509 (1989).

4. V. G. Knizhnik, A. M. Polyakov, and A. B. Zamolod-chikov, Mod. Phys. Lett. A **3**, 819 (1988).

5. A. M. Polyakov and A. B. Zamolodchikov, Mod. Phys. Lett. A **3**, 1213 (1988).

6. D. R. Karakhanyan, R. P. Manvelyan, and R. L. Mkrtchyan, Phys. Lett. B **329**, 185 (1994).

7. D. R. Karakhanyan, Phys. Lett. B **365**, 53 (1996).

8. S. Lazzarini and R. Stora, in *Knots, Topology and Quantum Field Theory: 13th John Hopkins Workshop*, Ed. by L. Lusanna (World Scientific, Singapore, 1989).

9. D. R. Karakhanyan, Yad. Fiz. **56** (9), 249 (1993) [Phys. At. Nucl. **56**, 1294 (1993)].

10. D. R. Karakhanian and A. G. Sedrakian, Phys. Lett. B **236**, 140 (1990); Phys. Lett. B **260**, 53 (1991).

11. T. Fujiwara, H. Igarashi, and T. Suzuki, Ann. Phys. **254**, 233 (1997).

12. M. Grisaru and R. Xu, Phys. Lett. B **205**, 486 (1988).

13. D. R. Karakhanyan, R. P. Manvelyan, and R. L. Mkrtchyan, Mod. Phys. Lett. A **11**, 409 (1996).

14. T. Fujiwara, Y. Igarashi, M. Koseki, *et al.*, Nucl. Phys. B **425**, 289 (1994).

15. K. Kamimura, J. Gomis, and R. Kuriki, Nucl. Phys. B **471**, 513 (1996).

Translated by R. Tyapaev

Reentrant Violation of Special Relativity in the Low-Energy Corner[1]

G. E. Volovik

Low Temperature Laboratory, Helsinki University of Technology, FIN-02015 HUT, Finland
Landau Institute for Theoretical Physics, Russian Academy of Sciences, ul. Kosygina 2, Moscow, 117334 Russia
Received January 25, 2001

In the effective relativistic quantum field theories, the energy region in which special relativity holds can be sandwiched from both the high- and low-energy sides by domains where special relativity is violated. An example is provided by ^{3}He-A, where the relativistic quantum field theory emerges as the effective theory. The reentrant violation of special relativity in the ultralow-energy corner is accompanied by the redistribution of the momentum-space topological charges among the fermionic flavors. At this ultralow energy, an exotic massless fermion with topological charge $N_3 = 2$ arises whose energy spectrum mixes classical and relativistic behaviors. This effect can lead to neutrino oscillations, if neutrino flavors are still massless on this energy scale. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 11.30.Cp; 14.60.Pq; 67.57.-z

INTRODUCTION

The condensed-matter analogy supports the idea that special and general relativity might be emergent properties of quantum vacuum that arise gradually in the low-energy corner [1–3]. If this is true, one can expect that the Lorentz invariance of our low-energy world is violated at high energy. The condensed matter provides examples of how this violation can occur. Here we demonstrate one generic example which is realized in the superfluid ^{3}He-A, where the effective special relativity and gravity arise in the low-energy corner together with chiral fermions and effective gauge fields [3]. It suggests that, if the effective Lorentz invariance is violated in the extreme limit of the "Planckian" scale, it becomes violated also in the opposite extreme limit of ultralow energy. The energy scale where the reentrant violation of special relativity occurs is also dictated by "trans-Planckian" physics. If there are still fermions which remain chiral and massless when approaching this ultralow-energy scale, the reentrant violation of Lorentz invariance leads to the crucial reconstruction of their energy spectrum. Thus, trans-Planckian physics can be probed in the limit of low energies.

In this example, two flavors of chiral left-handed fermions are hybridized to produce one massive fermion with the relativistic spectrum $E^2 = c^2p^2 + m^2c^4$ and one exotic gapless fermion whose spectrum mixes classical and relativistic behaviors: $E^2 = c^2 p_\parallel^2 + (p_\perp^2/2m)^2$ (see

figure). Such an energy spectrum is a consequence of the nontrivial momentum-space topology. The hybridization of fermions due to the violation of special relativity can provide a scenario for neutrino oscillations similar to that discussed in [4], where the violation of special relativity was also considered, though in terms of different maximum attainable velocities c for different neutrino species.

MOMENTUM-SPACE TOPOLOGY AND DISCRETE SYMMETRY BETWEEN FERMIONS

Special relativity (and also general relativity, with the effective gravitational field being one of the collective modes of the fermionic vacuum) naturally arises in those Fermi superfluids whose fermionic quasiparticle spectrum contains topologically nontrivial point nodes in the momentum space. Examples are the superfluid phases of ^{3}He: ^{3}He-A and the planar state. In the low-energy limit, i.e., in the vicinity of a given topologically stable point node (Fermi point), fermionic quasiparticles behave as chiral fermions with the massless spectrum obeying the relativistic-like equation

$$g_{(a)}^{\mu\nu}(p_\mu - p_{\mu(a)})(p_\nu - p_{\nu(a)}) = 0. \tag{1}$$

Here, $p_{\mu(a)}$ (position of the node in the spectrum of the ath quasiparticle) and $g_{(a)}^{\mu\nu}$ are the dynamic variables describing the collective bosonic degrees of freedom of the vacuum. They play the role of the gauge and gravity fields, respectively.

---

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

In each of the two phases of superfluid ^{3}He (^{3}He-A and the planar state), there is a symmetry that connects all the low-energy fermionic species. As a result, the effective metric $g_{\mu\nu}$ is the same for all fermions (at least in equilibrium), which means that all of them have the same "speed of light" (i.e., the same maximum attainable speed). Moreover, for the "perfect" fermionic system (for which the Lagrangian for the collective bosonic modes is obtained by the integration over fermions in the vicinity of the Fermi points, see [3]), the bosonic fields are governed by the same effective metric $g_{\mu\nu}$ as are the fermions and, hence, have the same speed of light. Thus, the nontrivial momentum-space topology and the symmetry between fermions are two ingredients for establishing special relativity in the low-energy corner of the effective theory.

The massless (gapless) character of the fermionic spectrum in a system with Fermi points is protected by the nonzero value of the ground-state topological invariant that is expressed as the integral over the Green's function in the 4D momentum–frequency space:

$$N_3 = \frac{1}{24\pi^2} e_{\mu\nu\lambda\gamma} \mathbf{tr} \int_\sigma dS^\gamma \mathcal{G} \partial_{p_\mu} \mathcal{G}^{-1} \mathcal{G} \partial_{p_\nu} \mathcal{G}^{-1} \mathcal{G} \partial_{p_\lambda} \mathcal{G}^{-1}. \quad (2)$$

The integral is over the surface σ enclosing Fermi point $p_{\mu(a)} = (\mathbf{p}_a, 0)$ in the 4D momentum space $p_\mu = (\mathbf{p}, p_0)$, p_0 is the energy (frequency) along the imaginary axis, and $\mathbf{tr}$ is the trace over the fermionic indices. If the topological charge of the Fermi point is $N_3 = +1$ or $N_3 = -1$, then in the vicinity of this point the quasiparticle is a massless fermion, whose Green's function, after proper rescaling and shifting the position of the Fermi point, has the following form

$$\mathcal{G} = (ip_0 - \mathcal{H})^{-1}, \quad \mathcal{H} = N_3 c \boldsymbol{\sigma} \cdot \mathbf{p}. \quad (3)$$

This is the Green's function of a left-handed (if $N_3 = -1$) or right-handed (if $N_3 = +1$) chiral fermion. Thus, the topologically nontrivial momentum-space topology automatically produces "relativistic" chiral fermions as low-energy quasiparticles if $N_3 = \pm 1$.

In the gapless superfluid phases of ^{3}He, the topological invariants of each of the two Fermi points are different from ± 1; they are $N_3 = \pm 2$ in ^{3}He-A and $N_3 = 0$ in the planar state. In both phases, however, there is a discrete symmetry between the fermionic species, which leads to the equal (in ^{3}He-A) or opposite (in the planar state) distribution of topological charges among fermions: $N_3 = +2 \longrightarrow +1 + 1$ and $N_3 = -2 \longrightarrow -1 - 1$ in ^{3}He-A, and $N_3 = 0 \longrightarrow +1 - 1$ in the planar state. As a result, each gapless fermion has a unitary charge, $N_3 = -1$ or $N_3 = +1$, and, hence, all gapless quasiparticles are relativistic in the low-energy corner. This again shows that the existence of a discrete or continuous symmetry between fermions is important for special relativity to hold in the low-energy corner. For a discrete symmetry
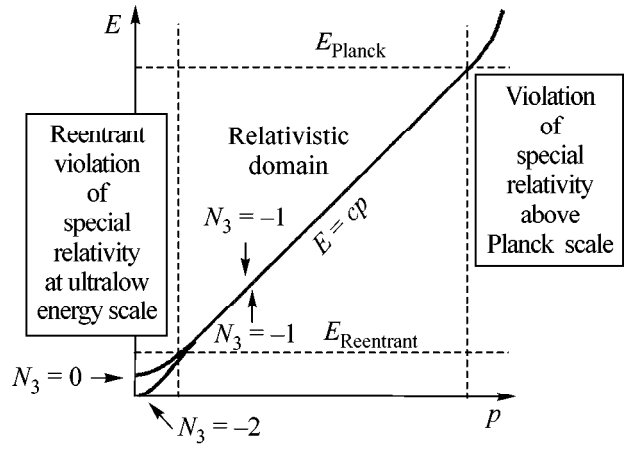
Low-energy memory of high-energy nonsymmetric physics.

providing unitary charges $N_3 = -1$ and $N_3 = +1$ for chiral fermions in the standard model, see [3].

CONDENSED MATTER SCENARIO OF THE REENTRANT VIOLATION OF SPECIAL RELATIVITY

In ^{3}He-A, the global symmetry $SO(3)_S \times SO(3)_L \times U(1)_N$ of the normal ^{3}He is broken into the little group $U(1)_{L_z - N/2} \times U(1)_{S_z}$ whose generators are $L_z - N/2$ and S_z. Here, $SO(3)_S$, $SO(3)_L$, and $U(1)_N$ are, correspondingly, the spin rotation group, the group of orthogonal coordinate transformations, and the global $U(1)$ group responsible for the conservation of the global charge—the number of ^{3}He atoms. The corresponding 3×3 order-parameter matrix $A_{\mu i}$, which transforms as a vector under spin rotations $SO(3)_S$ (the first index) and as a vector under orbital rotations $SO(3)_L$ (the second index), is

$$A_{\mu i} = \Delta_0 \hat{z}_\mu (\hat{x}_i + i\hat{y}_i). \quad (4)$$

Fermionic quasiparticles living in the vacuum with such order parameter have two point nodes in the spectrum. In the vicinity of the Fermi point at $\mathbf{p} = (0, 0, p_F)$ with the topological charge $N_3 = -2$, these quasiparticles correspond to two chiral left-handed relativistic fermions described by the following Bogoliubov–Nambu Hamiltonian

$$\mathcal{H}_{A-\text{phase}} = c_\parallel (p_z - p_F)\check{\tau}_3 + c_\perp \sigma_z (\check{\tau}_1 p_x - \check{\tau}_2 p_y). \quad (5)$$

Here, the speeds of light propagating along and perpendicular to the z axis are, correspondingly, $c_\parallel = v_F$ and $c_\perp = \Delta_0/p_F$; v_F and p_F are, respectively, the Fermi velocity and the Fermi momentum in the normal ^{3}He; Δ_0 is the amplitude of the gap; σ_z is the Pauli matrix for the nuclear spin of the ^{3}He atom; and $\check{\tau}_i$ are the Pauli

matrices for the Bogoliubov–Nambu spin in the particle–hole space.

Two projections of atomic spin $(1/2)\sigma_z = \pm 1/2$ can be considered as two fermionic flavors. Each of the two fermions has $N_3 = -1$, so that both fermions are chiral (left-handed) and massless with the energy spectrum

$$E^2 = c_\parallel^2 \tilde{p}_z^2 + c_\perp^2 p_\perp^2, \quad N_3 = -1, \tag{6}$$

where $\tilde{p}_z = p_z - p_F$. The symmetry, which couples the two flavors and forces them to have identical topological charge $N_3 = -1$ and identical relativistic spectrum, is the discrete symmetry U_2 of the order parameter in Eq. (4). U_2 is a combined symmetry; it is an element of the $SO(3)_S$ group, say, the spin π rotation about x axis supplemented by the gauge rotation $e^{i\pi}$ from the $U(1)_N$ group: $U_2 = C_\pi^x e^{i\pi}$. It is the same discrete symmetry as that giving rise to the Alice string (half-quantum vortex, see [5] for review). Such symmetry came from the fundamental level well above the first Planck scale $E_{Planck} = \Delta_0^2/v_F p_F$ [3] on which the spectrum becomes nonlinear and the Lorentz invariance is violated.

However, this discrete U_2 symmetry is not exact in ^{3}He even on the fundamental level. This is because of a tiny spin–orbit interaction, which slightly violates the $SO(3)_S$ symmetry under the separate rotations in spin space. Since the U_2 symmetry was instrumental for establishing special relativity in the low-energy corner, its violation must lead to the violation of the Lorentz invariance and also to the mixing of the two fermionic flavors at a very low energy determined by the tiny spin–orbit coupling.

Let us consider how this happens in ^{3}He-A. Due to the spin–orbit coupling, the $U(1)_{L_z - N/2} \times U(1)_{S_z}$ symmetry of ^{3}He-A is not exact. The remaining exact symmetry is a combined symmetry constructed from the sum of two generators: $U(1)_{J_z - N/2}$, where $J_z = S_z + L_z$ is the generator of the simultaneous rotations of spins and orbital degrees of freedom. As a result, the order parameter in Eq. (4) acquires a small correction consistent with the $U(1)_{J_z - N/2}$ symmetry:

$$A_{\mu i} = \Delta_0 \hat{z}_\mu (\hat{x}_i + i\hat{y}_i) + \alpha \Delta_0 (\hat{x}_\mu + i\hat{y}_\mu)\hat{z}_i. \tag{7}$$

The first term corresponds to the Cooper-pair state with $L_z = 1/2$ per atom and $S_z = 0$, while the second one is a small admixture of the state with $S_z = 1/2$ and $L_z = 0$. Both components have $J_z = 1/2$ and, thus, must be present in the order parameter. Due to the second term, the order parameter is not symmetric under the U_2 operation. The small parameter $\alpha \sim \xi^2/\xi_D^2 \sim 10^{-5}$ is the relative strength of spin–orbit coupling, where $\xi \sim 10^{-6}$–10^{-5} cm is the superfluid coherence length and $\xi_D \sim 10^{-3}$ cm is the so-called dipole length characterizing the spin–orbit coupling.

The Bogoliubov–Nambu Hamiltonian for fermionic quasiparticles in such a vacuum is modified compared to that in the pure vacuum state with $L_z = 1/2$ and $S_z = 0$ in Eq. (5):

$$\begin{aligned}\mathcal{H}_{A-phase} &= c_\parallel \tilde{p}_z \check{\tau}_3 + c_\perp \sigma_z (\check{\tau}_1 p_x - \check{\tau}_2 p_y) \\ &\quad + \alpha c_\perp p_z (\sigma_x \check{\tau}_1 - \sigma_y \check{\tau}_2).\end{aligned} \tag{8}$$

Diagonalization of this Hamiltonian shows that the small correction due to spin–orbit coupling gives rise to the following splitting of the energy spectrum

$$E_\pm^2 = c_\parallel^2 \tilde{p}_z^2 + c_\perp^2 \left(\alpha|p_z| \pm \sqrt{\alpha^2 p_z^2 + p_\perp^2}\right)^2. \tag{9}$$

Near the Fermi point one can set $\alpha|p_z| = \alpha p_F$. The $+$ and $-$ branches give the gapped and gapless spectra correspondingly. For $p_\perp \ll \alpha p_F$, one has

$$E_+^2 \approx c_\parallel^2 \tilde{p}_z^2 + \tilde{c}_\perp^2 p_\perp^2 + \tilde{m}^2 \tilde{c}_\perp^4, \tag{10}$$

$$E_-^2 \approx c_\parallel^2 \tilde{p}_z^2 + \frac{p_\perp^4}{4\tilde{m}^2}, \tag{11}$$

$$\tilde{c}_\perp = \sqrt{2c_\perp}, \quad \tilde{m} = \alpha\frac{p_F}{c_\perp}, \quad p_\perp \ll \tilde{m}c_\perp. \tag{12}$$

In this ultralow-energy corner, the gapped branch of the spectrum in Eq. (10) is relativistic, though with the speed of light other than in the intermediate regime of Eq. (6). The gapless branch in Eq. (11) is relativistic in one direction $E = c_\parallel|\tilde{p}_z|$, and is classical $E = p_\perp^2/2\tilde{m}$ for the motion in the transverse direction.

MOMENTUM-SPACE TOPOLOGY OF EXOTIC FERMION

What is important is that such a splitting of the spectrum is generic and can thus occur in other effective field theories such as the standard model. This is due to the topological properties of the spectrum: the mixing of the two fermionic flavors occurs with the redistribution of the topological charge among two fermions. In the relativistic domain, each of the two fermions has a topological charge of $N_3 = -1$. It is easy to check that in the ultralow-energy corner this is not the case. While the total topological charge of the Fermi point $N_3 = -2$ must be conserved, it is now redistributed among the fermions in the following manner: the massive fermion (with energy E_+) acquires the trivial topological charge $N_3 = 0$ (that is why it becomes massive), while another (with energy E_-) has the double topological charge $N_3 = -2$ (see figure). It is important that the topological charge $N_3 = -2$ describes a single fermionic species: it cannot split into two fermions with $N_3 = -1$ each. This exotic fermion with $N_3 = -2$ is gapless because of the nonzero value of the topological charge, but the energy spectrum of this fermion is nonlinear. Such a spectrum cannot be described in relativistic language.

Much as the $N_3 = \pm 1$ fermions are necessarily relativistic and chiral in the low-energy corner, the fermions with higher $|N_3|$ are necessarily nonrelativistic. The momentum-space topology, which induces special relativity if $|N_3| = 1$, becomes incompatible with the relativistic invariance if $|N_3| > 1$, and the latter is obligatory violated. The properties of the fermionic systems with multiple zeroes, $|N_3| > 1$, including the axial anomaly in its nonrelativistic version, are discussed in [6].

The energy scale on which the splitting of the energy spectrum occurs is $E_{\text{Reentrant}} = \alpha \Delta_0$, which is much less than the first Planck level in ^{3}He-A, $E_{\text{Planck}} = \Delta_0^2 / v_F p_F$. Thus, the relativistic region for the ^{3}He-A fermions, $E_{\text{Reentrant}} \ll E \ll E_{\text{Planck}}$, is sandwiched from both high and low energies by the nonrelativistic regions.

DISCUSSION

The above example of ^{3}He-A shows that the discrete symmetry between fermions, together with the momentum-space topology, guarantee that massless fermions obey special relativity in the low-energy corner. If the discrete symmetry is approximate, then in the ultralow-energy corner the redistribution of the momentum-space topological charges among the fermions occurs with appearance of a higher topological charge $|N_3| > 1$. This topological transition leads to a strong modification of the energy spectrum, which becomes essentially nonrelativistic.

In principle, such a topological transition with the appearance of exotic fermions with $N_3 = \pm 2$ can also occur in the relativistic quantum field theories if these theories are effective. In the effective theory, the Lorentz invariance (and hence, special relativity) appears in the low-energy corner as an emergent phenomenon, while it can be violated at a high energy approaching the Planck scale. At low energy, fermions are chiral and relativistic if there is a symmetry between the flavors. If such symmetry is violated, either spontaneously or due to the fundamental physics above the Planck scale, then in the extreme low-energy limit, when the asymmetry between the fermionic flavors

becomes important, the system remembers its high-energy nonrelativistic origin. The rearrangement of the topological charges N_3 among the fermionic species occurs and special relativity disappears again.

This scenario can be applied to massless neutrinos. The violation of the horizontal symmetry between the left-handed neutrino flavors can lead to the violation of Lorentz invariance at very low energy. If neutrinos remain massless at this ultralow-energy scale, then below this scale the two flavors (say, electronic and muonic left-handed neutrinos each with $N_3 = -1$) hybridize and produce an $N_3 = 0$ fermion with a gap and an exotic gapless $N_3 = -2$ fermion with the essentially nonlinear nonrelativistic spectrum. This is another example of the violation of special relativity, which can also give rise to the neutrino oscillations. The previously considered effect of the violation of special relativity on neutrino oscillations was related to the different speeds of light for different neutrino flavors [4] (the related effect is the violation of the weak equivalence principle: different flavors are differently coupled to gravity [7]).

REFERENCES

1. C. D. Frogatt and H. B. Nielsen, *Origin of Symmetry* (World Scientific, Singapore, 1991).
2. S. Chadha and H. B. Nielsen, Nucl. Phys. B **217**, 125 (1983).
3. G. E. Volovik, *Superfluid Analogies of Cosmological Phenomena*, Phys. Rep. (in press); gr-qc/0005091.
4. S. Coleman and S. L. Glashow, Phys. Lett. B **405**, 249 (1997); S. L. Glashow, A. Halprin, P. I. Krastev, *et al.*, Phys. Rev. D **56**, 2433 (1997).
5. G. E. Volovik, Proc. Natl. Acad. Sci. USA **97**, 2431 (2000).
6. G. E. Volovik and V. A. Konyshev, Pis'ma Zh. Éksp. Teor. Fiz. **47**, 207 (1988) [JETP Lett. **47**, 250 (1988)].
7. D. Majumdar, A. Raychaudhuri, and A. Sil, hep-ph/0009339; A. M. Gago, H. Nunokawa, and R. Zukanovich Funchal, hep-ph/0012168.

A Nonperturbative *a*-Expansion Technique and the Adler *D*-function[1]

A. N. Sissakian, I. L. Solovtsov, and O. P. Solovtsova

Joint Institute for Nuclear Research, Dubna, Moscow region, 141980 Russia

Received December 29, 2000

We show that the "experimental" *D*-functions corresponding to the e^+e^- annihilation into hadrons and the inclusive τ-decay data both are in good agreement with results obtained in the framework of the nonperturbative *a*-expansion method. © *2001 MAIK "Nauka/Interperiodica"*.

PACS numbers: 12.38.Lg; 13.65.+i; 13.35.Dx

The specific feature of quantum field theory is that a sufficiently complete study of the structure of a quantum field model within the framework of perturbative approach is insufficient even in theories with a small coupling constant. Numerous publications are devoted to the problem of going beyond the perturbation theory. However, many of them use model assumptions and phenomenological parameters which are not involved in the Lagrangian. Clearly, it is desirable to use a theoretical method that is based on a minimum number of additional parameters and allows the nonperturbative region to be considered. The theoretical method we use is the nonperturbative expansion technique [1] based on the idea of variational perturbation theory (see [2] for a review), which, in the case of QCD, leads to a new small expansion parameter *a*. Even when entering the infrared region of small momenta where the running coupling becomes large and the standard perturbative expansion fails, the *a*-expansion parameter remains small and the approach is still valid [3].

When comparing theoretical predictions with experimental data, it is important to relate measured quantities to the "simplest" theoretical objects, to check direct consequences of the theory without using model assumptions in an essential manner. Some single-argument functions, which are directly related to the experimentally measured quantities, can play the role of these objects. A theoretical description of inclusive processes can be made in terms of the functions of this sort. Let us mention, among them, the hadronic correlator $\Pi(s)$ and the corresponding Adler function [4] *D* that appear in the process of e^+e^- annihilation into hadrons and the inclusive decay of the τ lepton.

The cross section for e^+e^- annihilation into hadrons or its ratio to the leptonic cross section, $R(s)$, has a resonance structure, which is difficult to describe without model considerations at the present stage of theory. Moreover, the basic method of calculations in quantum field theory—perturbation theory—becomes ill-defined due to the so-called threshold singularities. These problems can, in principle, be avoided if one considers a smeared quantity [5]

$$R_\Delta(s) = \frac{\Delta}{\pi} \int_0^\infty ds' \frac{R(s')}{(s-s')^2 + \Delta^2}. \qquad (1)$$

However, the straightforward use of conventional perturbation theory for calculating R_Δ is impossible. Indeed, if the QCD contribution to the function $R(s)$ in Eq. (1) is, as usual, parametrized by the perturbative running coupling which has nonphysical singularities, it is difficult to define the integral on the right-hand side. Moreover, the standard renormalization group method gives a Q^2-evolution law for the running coupling in the Euclidean region, and the question arises of how to parametrize a quantity, e.g., $R(s)$ defined for timelike momentum transfers [6]. To perform this procedure self-consistently, it is important to maintain correct analytic properties of the hadronic correlator, which are violated in the perturbation theory. Within the nonperturbative *a* expansion, it is possible to maintain such analytic properties and to self-consistently determine the effective coupling in the Minkowskian region [7].[2]

Another function which characterizes the process of e^+e^- annihilation into hadrons and can be extracted

---

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

[2] The analytic approach to QCD [8] also leads to a well-defined procedure of analytic continuation [9].

from experimental data is the Adler function

$$D(Q^2) = Q^2 \int_0^\infty ds \frac{R(s)}{(s+Q^2)^2}. \qquad (2)$$

The *D*-function defined in the Euclidean region for a positive momentum Q^2 is a smooth function, and, thus, it is not necessary to apply any smearing procedure in order to compare theoretical results with experimental data. Recently, an "experimental" curve for this function has been obtained [10].

For massless quarks, one can write the Minkowskian quantity $R(s)$ in the form

$$R(s) = 3 \sum_f q_f^2 [1 + r_0 \lambda_s^{\text{eff}}(s)], \qquad (3)$$

where the sum runs over quark flavors, q_f are quark charges, and r_0 is the first perturbative coefficients that is independent of the renormalization scheme. This expression includes the effective coupling defined in the Minkowskian region or, as we will say, in the *s* channel, which is reflected in the subscript *s*. It should be stressed that, as it is argued from general principles, the behavior of the effective couplings in the spacelike and the timelike domains cannot be symmetric [11].

Within the *a*-expansion method, the *s*-channel running coupling can be written as

$$\lambda_s^{(i)}(s) = \frac{1}{2\pi i} \frac{1}{2\beta_0} [\phi^{(i)}(a_+) - \phi^{(i)}(a_-)], \qquad (4)$$

where $a_\pm$ obeys the equation [2]

$$F(a_\pm) = F(a_0) + \frac{2\beta_0}{C} \left(\ln \frac{s}{Q_0^2} \pm i\pi \right). \qquad (5)$$

At the $O(a^3)$ level, the function $\phi(a)$ has the form

$$\phi^{(3)}(a) = -4 \ln a - \frac{72}{11} \frac{1}{1-a}$$
$$+ \frac{318}{121} \ln(1-a) + \frac{256}{363} \ln \left(1 + \frac{9}{2}a \right). \qquad (6)$$

Similarly, a more complicated expression for the $O(a^5)$ level, which we will use, can be derived.

A convenient way to incorporate quark mass effects is to use an approximate expression [5]

$$\tilde{R}(s) = 3 \sum_f q_f^2 \Theta(s - 4m_f^2) \mathcal{R}_f(s),$$
$$\mathcal{R}_f(s) = T(v_f)[1 + g(v_f)r_f(s)], \qquad (7)$$

where

$$T(v) = v \frac{3-v^2}{2},$$

$$g(v) = \frac{4\pi}{3} \left[\frac{\pi}{2v} - \frac{3+v}{4} \left(\frac{\pi}{2} - \frac{3}{4\pi} \right) \right], \qquad (8)$$

$$v_f = \sqrt{1 - \frac{4m_f^2}{s}}.$$

The quantity $r_f(s)$ is defined by the *s*-channel effective coupling $\lambda_s^{\text{eff}}(s)$. Smeared quantity (1) and the *D* function (2) can be calculated by using Eq. (7) in the corresponding integrands. For MS-like renormalization schemes, one has to consider some matching procedure. To perform this matching procedure, we can require that the *s*-channel running coupling and its derivative be continuous functions in the vicinity of the threshold [7, 12].

A description of quark–antiquark systems near the threshold requires taking into account the resummation factor. In a nonrelativistic approximation, this is the well-known Sommerfeld–Sakharov factor [13]. For a systematic relativistic analysis of quark–antiquark systems, it is essential to have from the very beginning a relativistic generalization of this factor. A new form for this relativistic factor in the case of QCD was proposed in [14] by using the quasipotential approach to quantum field theory formulated in the relativistic configuration representation [15]. The local Coulomb potential defined in this representation is specified by its QCD-like behavior in momentum space [16].

The relativistic *S* factor has the form [14]

$$S(\chi) = \frac{X(\chi)}{1 - \exp[-X(\chi)]}, \quad X(\chi) = \frac{4\pi\alpha_S}{3\sinh\chi}, \qquad (9)$$

where χ is the rapidity, which is related to *s* by $2m\cosh\chi = \sqrt{s}$.

To take into account the threshold resummation factor, we modify Eq. (7) as follows:

$$\mathcal{R}(s) = T(v) \left[S(\chi) - \frac{1}{2}X(\chi) + g(v)r(s) \right]. \qquad (10)$$

As the mass $m \longrightarrow 0$, this expression leads to Eq. (3). We use Eq. (10) in our analysis.

The nonstrange vector contribution to the inclusive τ-lepton decay can be described by analogy with the process of e^+e^- annihilation into hadrons. Using the theoretical expression for the R_τ ratio [17]

$$R_\tau^V = R^{(0)} \int_0^{M_\tau^2} \frac{ds}{M_\tau^2} \left(1 - \frac{s}{M_\tau^2} \right)^2 \left(1 + \frac{2s}{M_\tau^2} \right) \mathcal{R}(s), \qquad (11)$$
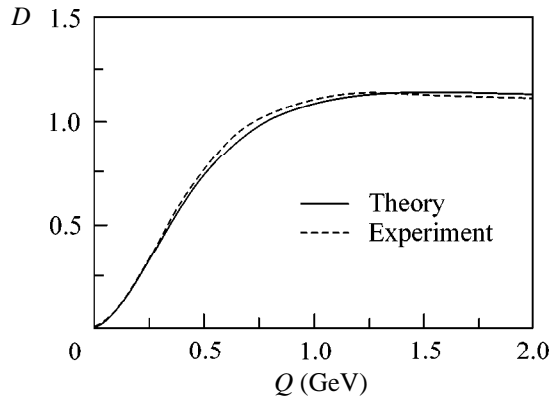
Fig. 1. Plot of the light D function. The experimental curve corresponding to the ALEPH data is taken from [19].
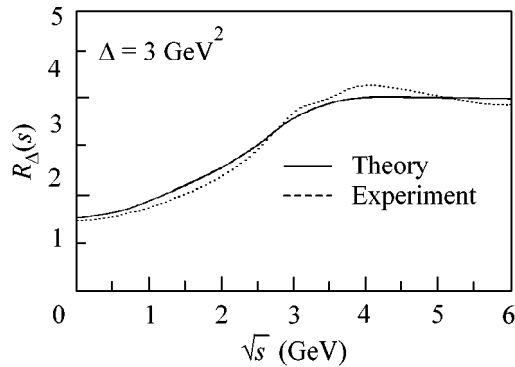

Fig. 2. The smeared quantity $R_\Delta(s)$ for $\Delta = 3$ GeV^2. The solid curve is our result. The smeared experimental curve is taken from [22].
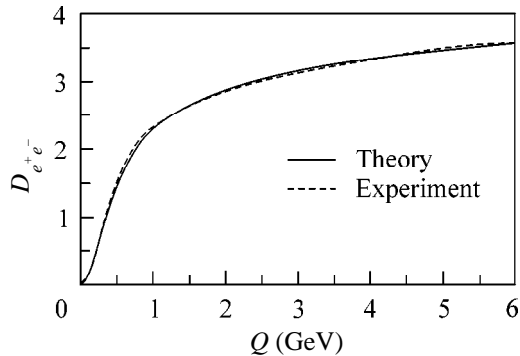

Fig. 3. The D function for the process of e^+e^- annihilation into hadrons. The solid curve is our result for five active quarks. The experimental curve is taken from [10].

where $R^{(0)}$ corresponds to the parton level, and the measured value $R_\tau^V = 1.775 \pm 0.017$ [18] is an input, we extracted the value of parameter a_0 in Eq. (5) at the τ-mass scale, $Q_0 = M_\tau$.

The "light" D function with three active quarks is shown in Fig. 1, where we draw the experimental curve (dashed line), which was extracted in [19] from the ALEPH data, and our theoretical result (solid line) obtained by using the following effective masses of light quarks: $m_u = m_d = 260$ MeV and $m_s = 400$ MeV.[3] These values are close to the constituent quark masses and incorporate some nonperturbative effects. The shape of the infrared tail of the D function is sensitive to the value of these masses.

In Fig. 2, we present the smeared function $R_\Delta(s)$ for $\Delta = 3$ GeV^2. We use the same masses for the light quarks as before and the following masses for heavy quarks $m_c = 1.3$ GeV and $m_b = 4.7$ GeV. The smeared $R_\Delta(s)$ function for $\Delta \simeq 1$–3 GeV^2 is less sensitive to the value of light quark masses, as compared with the infrared tail of the D function. The result for the D function of the e^+e^- annihilation process, which includes both light and heavy quarks, is plotted in Fig. 3. The experimental curve is taken from [10].

The experimental D function turned out to be a smooth function without traces of the resonance structure of $R(s)$. One can expect that this object more precisely reflects the quark–hadron duality and is convenient for comparing theoretical predictions with experimental data. Note here that any finite order of the operator product expansion fails to describe the infrared tail of the D function. Within the framework of nonperturbative a-expansion technique, we have obtained a good agreement between our results and the experimental data down to the lowest energy scale for both Minkowskian and Euclidean quantities.

REFERENCES

1. I. L. Solovtsov, Phys. Lett. B **327**, 335 (1994); **340**, 245 (1994).

2. A. N. Sissakian and I. L. Solovtsov, Fiz. Élem. Chastits At. Yadra **25**, 1127 (1994) [Phys. Part. Nucl. **25**, 478 (1994)]; Fiz. Élem. Chastits At. Yadra **30**, 1057 (1999) [Phys. Part. Nucl. **30**, 461 (1999)].

3. H. F. Jones and I. L. Solovtsov, Phys. Lett. B **349**, 519 (1995); in *Proceedings of the International European Conference on High Energy Physics, Brussels, 1995*, Ed. by J. Lemonne, C. Vander Velde, and F. Verbeure (World Scientific, Singapore, 1996), p. 242.

4. S. L. Adler, Phys. Rev. D **10**, 3714 (1974).

5. E. C. Poggio, H. R. Quinn, and S. Weinberg, Phys. Rev. D **13**, 1958 (1976).

6. A. V. Radyushkin, Preprint No. E2-82-159, JINR (Joint Institute for Nuclear Research, Dubna, 1982); JINR

---

[3] Virtually the same values were used to describe the experimental data in [20, 21].

Rapid Commun. **4**, 9 (1996); N. V. Krasnikov and A. A. Pivovarov, Phys. Lett. B **116B**, 168 (1982).

7. H. F. Jones, I. L. Solovtsov, and O. P. Solovtsova, Phys. Lett. B **357**, 441 (1995).

8. D. V. Shirkov and I. L. Solovtsov, JINR Rapid Commun. **2**, 5 (1996); Phys. Rev. Lett. **79**, 1209 (1997).

9. K. A. Milton and I. L. Solovtsov, Phys. Rev. D **55**, 5295 (1997); K. A. Milton and O. P. Solovtsova, Phys. Rev. D **57**, 5402 (1998).

10. S. Eidelman, F. Jegerlehner, A. L. Kataev, and O. Veretin, Phys. Lett. B **454**, 369 (1999).

11. K. A. Milton and I. L. Solovtsov, Phys. Rev. D **59**, 107701 (1999).

12. K. A. Milton, I. L. Solovtsov, and O. P. Solovtsova, Eur. Phys. J. C **13**, 497 (2000).

13. A. Sommerfeld, in *Atombau und Spektrallinien* (Vieweg, Braunschweig, 1939), Vol. II; A. D. Sakharov, Zh. Éksp. Teor. Fiz. **18**, 631 (1948).

14. K. A. Milton and I. L. Solovtsov, hep-ph/0005175.

15. V. G. Kadyshevsky, R. M. Mir-Kasimov, and N. B. Skachkov, Nuovo Cimento A **55**, 233 (1968); Fiz. Élem. Chastits At. Yadra **2**, 635 (1972) [Sov. J. Part. Nucl. **2**, 69 (1972)].

16. V. I. Savrin and N. B. Skachkov, Lett. Nuovo Cimento **29**, 363 (1980).

17. E. Braaten, S. Narison, and A. Pich, Nucl. Phys. B **373**, 581 (1992).

18. ALEPH Collaboration (R. Barate, D. Buskulic, D. Decamp, *et al.*), Eur. Phys. J. C **4**, 409 (1998).

19. S. Peris, M. Perrottet, and E. de Rafael, JHEP **05**, 011 (1998); hep-ph/9805442.

20. A. I. Sanda, Phys. Rev. Lett. **42**, 1658 (1979).

21. D. V. Shirkov and I. L. Solovtsov, hep-ph/9906495.

22. A. C. Mattingly and P. M. Stevenson, Phys. Rev. D **49**, 437 (1994); S. J. Brodsky, J. R. Peláez, and N. Toumbas, Phys. Rev. D **60**, 037501 (1999).

Focusing of Nonlinear Wave Groups in Deep Water[1]

C. Kharif, E. Pelinovsky*, T. Talipova*, and A. Slunyaev*

Institut de Recherche sur les Phenomenes Hors Equilibre, Parc Scientifique et Technologique de Luminy,
13288 Marseille, Cedex 9, France

e-mail: kharif@pollux.univ-mrs.fr

* Laboratory of Hydrophysics and Nonlinear Acoustics, Institute of Applied Physics,
Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603600 Russia

e-mail: enpeli@hydro.appl.sci-nnov.ru

Received November 8, 2000; in final form, January 15, 2001

The freak wave phenomenon in the ocean is explained by the nonlinear dynamics of phase-modulated wave trains. It is shown that the preliminary quadratic phase modulation of wave packets leads to a significant amplification of the usual modulation (Benjamin–Feir) instability. Physically, the phase modulation of water waves may be due to a variable wind in storm areas. The well-known breather solutions of the cubic Schrödinger equation appear on the final stage of the nonlinear dynamics of wave packets when the phase modulation becomes more uniform. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 92.10.Hm; 47.35.+i; 92.60.Dj

The increase in a number of reported damages of ships and offshore platforms is explained very often by a freak wave appearance on the sea surface [1, 2]. Several physical mechanisms of the freak wave phenomenon are discussed. First of all, the water wave interaction with an opposite current is considered as a mechanism of wave amplification due to the blocking of water waves on the current. This phenomenon is investigated within the framework of the wave action balance equation and the variable-coefficient nonlinear Schrödinger equation [2–4]. The second mechanism of wave amplification is related to the formation of caustics in the wave field on random currents [5]. These theories are used to explain freak wave formation due to the Agulhas current off the southeast coast of South Africa. Many observations of abnormal waves have been done in areas with no strong currents. For such areas, the opinion that the nonlinearity of surface waves in deep water can produce a giant wave by itself has become very popular [6, 7]. The theory is based on the modulation instability of water waves (see review [8]) and the existence of breatherlike solutions of the nonlinear Schrödinger equation [7, 9, 10–13]. The amplitude of breathers can exceed the amplitude of unperturbed nonmodulated wave trains more than twice (remember that it is the formal definition of a freak

wave). The nonlinear Schrödinger equation is a simplified model of real wind waves, and more sophisticated models are also applied (Zakharov equation, Dysthe equation, etc.). In [6], the freak wave formation due to a modulation instability computed by a numerical model of the full-nonlinear hydrodynamic potential equations was compared with the cubic nonlinear Schrödinger equation and found a good agreement if the steepness of waves is not too large. According to these theories, freak waves exist during the characteristic time scale of modulation instability and may propagate at a relatively large distance. Meanwhile, the event descriptions emphasize the very short-lived character of the freak wave. In our opinion, the mechanism of water wave packet focusing due to phase (frequency) modulation should play a significant role in the formation of a short-lived freak wave. This mechanism is well known in the linear theory of dispersive waves [14] and may occur under specific meteorological conditions. For instance, an increase in wind speed generates, at the early stages, wave packets with low group velocities and, later, wave packets with larger group velocities. The propagation process results in the formation of an impulse of very large amplitude, which is due to the superposition of many spectral packets. Analytic solutions proving this linear focusing mechanism are presented in [15]. The phenomenon of significant wave focusing in laboratory tanks was reported [16, 17] for a wide variation of the wavelength/depth ratio for deep water and shallow water as well. Recently, the authors of [18] have

---

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

showed that the mechanism of wave focusing can be applied in the weakly nonlinear theory of shallow water[2] (Korteweg–de Vries model) and suggested a way to find possible forms of wave trains moving towards the freak wave, including the random background of wind waves. Owing to the absence of the modulation instability for shallow water, the wave focusing mechanism seems to be major in shallow water.

This paper deals with freak wave formation in deep water due to the focusing of nonlinear wave packets with phase modulation. This mechanism is compared with the possible generation of giant waves (breathers) due to the usual modulation instability of water waves. The main result of the paper is that the frequency modulation of a nonlinear wave field can lead to larger amplification of the freak wave than does the amplitude modulation usually considered by previous authors.

The simple model of weakly nonlinear deep-water wave packets is the famous cubic Schrödinger equation

$$i\frac{\partial A}{\partial t} + \frac{\partial^2 A}{\partial x^2} + 2|A|^2 A = 0, \qquad (1)$$

where, in dimensionless variables, A is proportional to the wave steepness, $A = \sqrt{2}\, k_0 a$, a is the amplitude of the surface elevation, k_0 and ω_0 are the carrier wave number and frequency respectively, $x = 2k_0 x' - \omega_0 t'$, $t = \omega_0 t'/2$, x' and t' are coordinate and time. Due to the invariance transformation of Eq. (1), $t \longrightarrow -t$, $i \longrightarrow -i$, the simplest algorithm to find the nonlinear wave packets moving toward the giant wave can be suggested: take the expected form of the freak wave as an initial condition for Eq. (1) and consider the resulting field as the initial condition that gives the freak wave under the invariant transformation. The solution of the Cauchy problem for the nonlinear Schrödinger equation is known by using the inverse scattering method, see the pioneering paper [21]. In general, the scattering data include both continuous and discrete spectra. The continuous spectrum corresponds to the dispersive wave packets. In the case of no discrete spectrum, the solution of Eq. (1) tends to the phase-modulated wave for large times

$$A(x, t) = \frac{Q}{\sqrt{t}} \exp\left[i\left(\frac{x^2}{4t} + 2Q^2 \ln t + \theta\right)\right], \qquad (2)$$

where Q and θ are functions of (x/t) [22]. When Q and θ are real constants, Eq. (2) gives an exact solution of Eq. (1), which almost coincides with the self-similar solution of the linear version of Eq. (1). The difference is in the logarithmic term of the phase. Replacing t by $T - t$ and i by $-i$, Eq. (2) describes the transformation of

the initial frequency modulated wave into the delta function formally representing the freak wave. Therefore, the mechanism of wave focusing is valid in linear theory and in nonlinear theory as well, but the nonlinearity influences the optimal phase (wavenumber) distribution of individual waves due to the logarithmic term that depends on wave intensity.

Each discrete eigenvalue λ of the spectrum corresponds to an isolated soliton with amplitude $2a$ and speed $4b$, where $a = \mathrm{Re}(\lambda)$ and $b = \mathrm{Im}(\lambda)$ [21]. The number of the discrete eigenvalues depends on the form and energy of the initial disturbance. If the expected freak wave has the form of an isolated crest with vanishing tails at infinity and no phase modulation, all discrete eigenvalues are real and correspond to standing solitary waves. Due to interaction between them, the resulting wave field is unsteady and shows the complex picture of the oscillating impulses of nonconstant form. In particular, the two-soliton solution (bi-soliton) describes a wave which breathes with the period $T = \pi/2\left|a_2^2 - a_1^2\right|$ and peak value $A_{\max} = 2(a_1 + a_2)$, where $a_1 = \mathrm{Re}(\lambda_1)$ and $a_2 = \mathrm{Re}(\lambda_2)$. The discrete spectrum is found for several forms of initial disturbances; in particular, [23] considered the profile $A_{fr}(x) = A_p \,\mathrm{sech}\,(x/L)$ (index fr refers to the freak wave). Eigenvalues are positive and equal to

$$\lambda_n L = \frac{M}{\pi} - n + \frac{1}{2}, \quad n = 1, 2, \ldots, N, \qquad (3)$$

where the number of eigenvalues is

$$N = E\left[\frac{M}{\pi} + \frac{1}{2}\right], \qquad (4)$$

E is the integer function and M is the mass of the freak wave, $M = \pi A_p L$.

It is important to note that if $M < \pi/2$, there is no soliton generation, and this case can be considered as quasi-linear. The wave evolves like a self-similar solution (2): at first, the wave focuses on short freak wave and then disperses. One soliton forms if $\pi/2 < M < 3\pi/2$, and its amplitude will vary from 0 to $2A_p$. In the latter case, the soliton amplitude exceeds the amplitude of the initial disturbance. If we consider such a wave group (one soliton with amplitude $2A_p$ and the dispersive tail) as an initial condition, it will transform into the sech-disturbance, but it has no specificity of the expected freak wave (its amplitude should be large on the background of other waves). If we introduce the formal definition of the freak wave (its amplitude is no less than twice the background amplitude), it means that the amplitude of the freak wave should satisfy the following condition

$$M < 2\pi/3. \qquad (5)$$

---

[2] The process of focusing of phase-modulated impulses in nonlinear media is also known and applied in optics [19, 20].
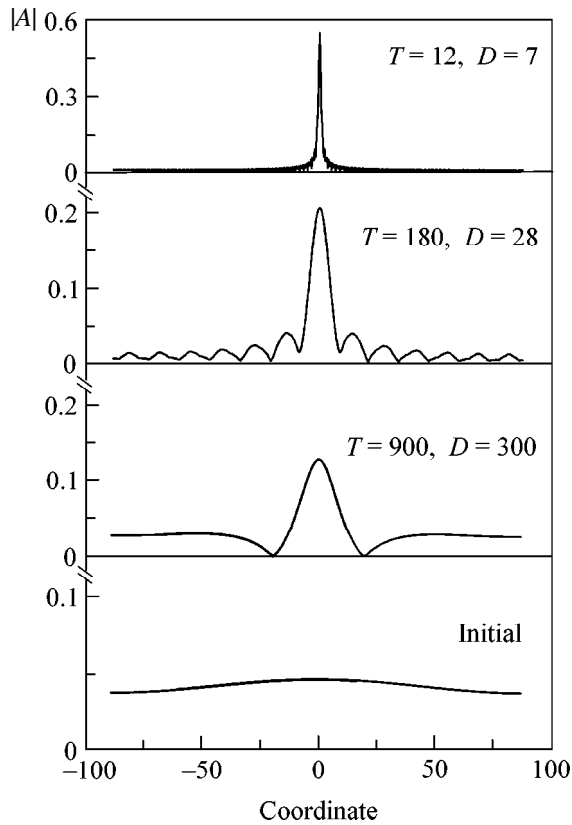
Fig. 1. Development of the modulation instability for different phase indexes *D*. Time of appearance of the wave of maximal amplitude is provided.

Therefore, the freak wave generated from the bounded wave group is a weakly nonlinear wave. Such a wave can be generated by dispersive wave packets only if $M < \pi/2$ or by dispersive wave packets plus the single soliton if $\pi/2 < M < 2\pi/3$. The interaction between solitons cannot generate the large-amplitude impulse; it will be comparable with solitons in amplitude. If we consider other profiles of the expected freak wave different from the sech-function, the integer constants in Eqs. (3) and (4) will change (see, for instance, [24]), but in the same order of magnitude. If we consider the initial impulse with the quadratic phase modulation like $\exp(iqx^2)$, the discrete eigenvalues will increase with q [25]. This result is obvious, because such a disturbance will transform first into an impulse with no phase modulation due to the wave focusing, and this large impulse leads to the large eigenvalues. Therefore, the form of the expected freak wave can be taken with no phase modulation, and the details of the waveform have no principal significance for the understanding of the wave focusing phenomenon in the nonlinear medium. The mechanism described above for the freak wave formation from the bounded wave packets is the same as for shallow water [18].

The mechanism of the localized wave formation (solitons or breathers) from the preliminary plane wave due to the modulation instability has been studied for almost 20 years. Several nonlinear structures can be considered as models for the freak wave, its peak amplitude exceeds more than twofold the unperturbed value. First of all, there is the Ma-breather [7, 9]

$$A(x, t)$$
$$= \frac{\cos(\Omega t - 2i\varphi) - \cosh(\varphi)\cosh(px)}{\cos(\Omega t) - \cosh(\varphi)\cosh(px)}\exp(2it), \quad (6)$$

where $p = 2\sinh(\varphi)$, $\Omega = 2\sinh(2\varphi)$, and φ is an arbitrary positive constant. This wave tends to the unperturbed plane wave of unit amplitude for $|x| \longrightarrow \infty$, and its amplitude is periodic in time with a frequency of Ω. The peak value of the breather exceeds the unperturbed value twofold and more for $0 < \varphi < 0.96$. It is important to emphasize that the freak wave phenomenon has a periodic character in this model.

Another solution, called homoclinic orbit, was found in [11, 13] and may be expressed by Eq. (6) if replacing φ by $-i\varphi$, p by $-ip$ and Ω by $-i\Omega$. This wave is periodic in space and tends to the unperturbed plane wave when $|t| \longrightarrow \infty$. The maximal peak value of the impulse is less than three. This homoclinic orbit can also be considered as a model for the freak wave for $\varphi > \pi/3$. It is important to mention that freak waves in this model should appear simultaneously in many spatial points.

Both breather solutions considered above for $\varphi = 0$ transform into the algebraic breather [10] as the peak value becomes three times higher than the value of the unperturbed wave amplitude. At the moment of maximal amplification ($t = 0$), the freak wave represents the large crest above the unperturbed plane wave ($|x| < 1/2$), and two depressions up to zero. The mass of this positive crest only is

$$M = \int_{-1/2}^{+1/2} |A(x, 0)|dx = 1 + 4\operatorname{arctanh}(1/2) \approx 2.9. \quad (7)$$

Thus, the breather solution provides a model for the freak wave with mass greater than for pure focusing regime, see Eq. (5). This is the main kinematic difference between focused and nonlinear freak waves. Physically, this difference can be clarified as follows. The focused freak wave is formed by the superposition of many spectral components and the number of spectral components, or the effective spectrum width, K_0, should be large to provide the narrow crest. The focused freak wave is a very weakly nonlinear dispersive wave, and it should be narrow if the dispersion prevails over nonlinearity. Its time of existence can be very small. The nonlinear freak wave is due to the modula-

tion instability. As it is well-known (see, for instance, [8]), the width of the unstable spectral domain is $K_{BF} = 2A_0$, and the characteristic time scale of the instability is $T_{BF} \sim 1/(2A_0^2)$, where A_0 is the amplitude of the unperturbed plane wave. For a nonlinear freak wave, dispersion and nonlinearity are of the same order. Its time of existence is the characteristic time of the modulation instability.

For small wave amplitudes, the width of the modulation instability is narrow and the focusing mechanism should dominate. With the increase of the unperturbed wave amplitudes, K_{BF} will become comparable to K_0, and the spectral components will contribute in both processes of the formation of the freak wave. In the case of no specific phase modulation of the wave packet, the nonlinear mechanism of the freak wave phenomenon should dominate the general dynamics of the wave field. But if the specific order of the "draw up" of the spectral components is organized (for instance due to the wind action), the phase modulation can cardinally change the modulation instability. The important role of frequency modulation in the modulation instability was emphasized in [26]. The author of [26] pointed out that the small focusing effects may have a destabilizing effect under certain conditions. Nevertheless, it seems that the nonlinear stage of the modulation instability for frequency-modulated wave packets was as yet not investigated in the literature. The effect of the quadratic phase modulation on the nonlinear evolution of the modulation instability is considered here numerically.

The nonlinear Schrödinger equation is solved by using a pseudospectral method in a periodic domain of dimensionless length 176; the number of points varies from 128 to 1024. The initial condition is

$$A(x, 0) = A_0(1 + 0.1\cos(x/d))\exp(ix^2/D^2), \quad (8)$$

where $A_0 = 0.043$, $d = 28$, and D is varied in wide range. Figure 1 shows various forms of the freak wave for different phase indexes D. The maximal amplitude is reached at different moments of time; they are given in Fig. 1. The phase modulation of the initial envelope leads to the increase in the wave amplitude and to the decrease in time of freak wave formation (the latter is provided in Fig. 2). For small D, the formation time is described by the power-law asymptotic expression ($T \sim D^2/4$), as it can be shown in the linear theory; for large D it tends to a constant value defined by the modulation instability (approximately $3T_{BF}$). As it is predicted, the phase modulation of the preliminary amplitude-modulated envelope leads to the formation of a more energetic wave impulse at a shorter time.

For $D = 28$, the long-term nonlinear dynamics of the wave field is displayed in Fig. 3. The phase modulation leads to a complex picture of a wave envelope with one
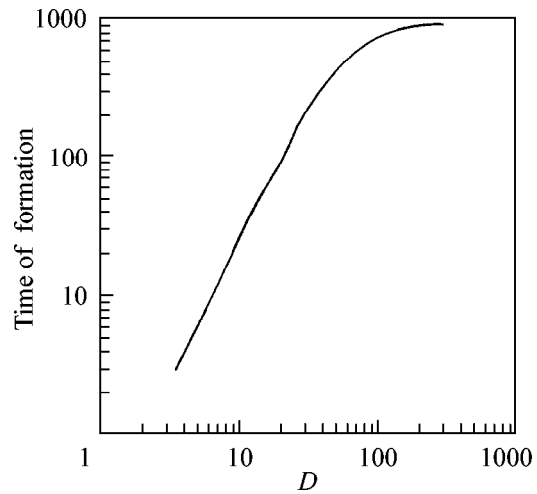
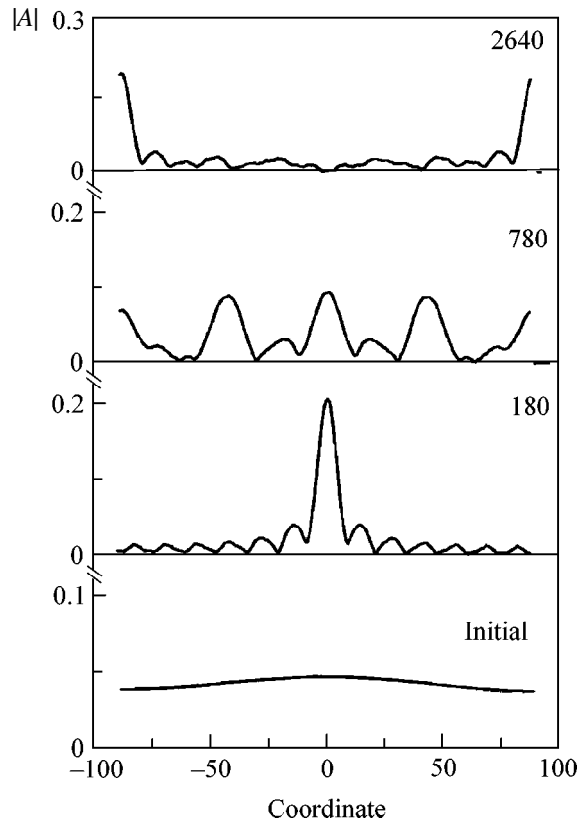**Fig. 2.** Time of appearance of the freak wave as a function of D.

Fig. 3. Wave evolution from the initial modulated disturbance ($D = 28$) at different times.

or several peaks and holes that can be considered as a group of freak waves. The time evolution of the maximal value of wave amplitude is shown in Fig. 4a. The very large amplitude peaks appear several times during 12000 time units, and their amplitudes decrease with time. Then the process becomes more stationary and
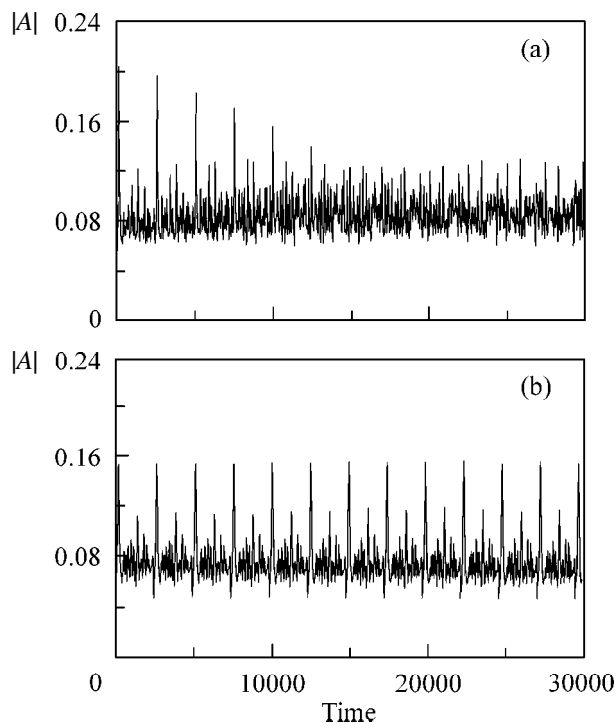
Fig. 4. Time variation of the maximal amplitude in the wave packets for $D = 28$; (a) nonlinear and (b) linear cases.

peaks with amplitudes about 0.12 appear regularly. Figure 4b shows the time evolution within the framework of linear theory. In this case, the process of generation of significant peaks is almost periodic. Similar behavior is found for wide ranges of the variation of the phase index D. Physically, the role of nonlinearity in the wave field behavior can be explained as follows. The quadratic phase modulation in Eq. (8) corresponds to a linear variation of the wave number (wave frequency) with distance. At the first stage, the slope of $K(x)$ increases and tends to infinity, leading to focusing (it can be easily shown for linear waves within the framework of the kinematic equation for the wave number). After focusing, the slope of $K(x)$ changes its sign and decreases. Then, a jump in wave number is formed and the function $K(x)$ becomes multivalued with many jumps. As a result, the wave packet in the periodic problem can focus many times, as it is shown in Fig. 4b. The nonlinear effects leads to the smoothing and uniformity of the phase distribution. Here, the role of the classical modulation instability is more significant and the wave transformation is similar to that studied in [6, 12, 13]. At this stage, the amplitude of freak wave is less than three times the amplitude of the unperturbed value.

Therefore, the effect of phase modulation of the initial wave disturbance leads to a significant intensification of the process of freak wave generation. The phase modulation of the wind wave field can be due to specific meteorological conditions and the relation

between the observed freak waves and heavy weather conditions is very often mentioned in the literature. Due to the short time of existence of the freak wave, the random forcing from the wind (this process should be studied within the framework of the forced version of the nonlinear Schrödinger equation) cannot modify radically the process of freak wave formation from the frequency-modulated disturbances at least at the first stage, meanwhile as it was shown in [27], the usual modulation instability is reduced in the random fields.

The one-dimensional model used cannot predict the behavior of the wave field in two-dimensional case. The transversal instability (see [28]) may have an influence on the process of formation of a freak wave. This should become the topic for following investigations.

REFERENCES

1. S. E. Sand, N. E. O. Hansen, P. Klinting, *et al.*, in *Water Wave Kinematics*, Ed. by O. Torum and O. T. Gudmestad (Kluwer, Dordrecht, 1990), p. 535.
2. I. Lavrenov, Nat. Hazards **17**, 117 (1998).
3. D. H. Peregrine, Adv. Appl. Mech. **16**, 9 (1976).
4. R. Smith, J. Fluid Mech. **77**, 417 (1976).
5. B. S. White and B. Fornberg, J. Fluid Mech. **355**, 113 (1998).
6. K. L. Henderson, D. H. Peregrine, and J. W. Dold, Wave Motion **29**, 341 (1999).
7. K. B. Dysthe and K. Trulsen, Phys. Scr. T **T82**, 48 (1999).
8. F. Dias and C. Kharif, Annu. Rev. Fluid Mech. **31**, 301 (1999).
9. Y. Ch. Ma, Stud. Appl. Math. **60**, 43 (1979).
10. D. H. Peregrine, J. Aust. Math. Soc. Ser. B, Appl. Math. **25**, 16 (1983).
11. N. N. Akhmediev, V. M. Eleonskii, and N. E. Kulagin, Zh. Éksp. Teor. Fiz. **89**, 1542 (1985) [Sov. Phys. JETP **62**, 894 (1985)].
12. B. M. Herbst and M. A. Ablowitz, Phys. Rev. Lett. **62**, 2065 (1989).
13. M. J. Ablowitz and B. M. Hersbst, SIAM J. Appl. Math. **50**, 339 (1990).
14. L. Ostrovsky and A. Potapov, *Modulated Waves. Theory and Applications* (John Hopkins Univ. Press, Baltimore, 1999).
15. E. Pelinovsky and C. Kharif, in *Proceedings of the 15th International Workshop on Water Waves and Floating Bodies, Israel, 2000*, Ed. by T. Miloh and G. Zilman, p. 142.
16. S. P. Kjeldsen, in *Water Wave Kinematics*, Ed. by O. Torum and O. T. Gudmestad (Kluwer, Dordrecht, 1990), p. 453.
17. M. G. Brown and A. Jensen, submitted to J. Geophys. Res.

18. E. Pelinovsky, T. Talipova, and C. Kharif, Physica D (Amsterdam) **147**, 83 (2000).

19. S. A. Akhmanov, V. A. Visloukh, and A. S. Chirkin, *Optics of Femtosecond Laser Impulses* (Nauka, Moscow, 1988).

20. A. P. Sukhorukov, in *Nonlinear Wave Interactions in Optics and Radiophysics* (Nauka, Moscow, 1988), p. 55.

21. V. E. Zakharov and A. B. Shabat, Zh. Éksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys. JETP **34**, 62 (1972)].

22. M. J. Ablowitz and H. Segur, J. Fluid Mech. **92**, 691 (1979).

23. J. Satsuma and N. Yajima, Suppl. Prog. Theor. Phys. **55**, 284 (1974).

24. S. Clarke, R. Grimshaw, P. Miller, *et al.*, Chaos **10**, 383 (2000).

25. D. J. Kaup and B. A. Malomed, Physica D (Amsterdam) **84**, 319 (1995).

26. C. A. van Duin, J. Fluid Mech. **399**, 237 (1999).

27. I. E. Alber, Proc. R. Soc. London, Ser. A **363**, 525 (1978).

28. V. E. Zakharov and A. M. Rubenchik, Zh. Éksp. Teor. Fiz. **65**, 997 (1973) [Sov. Phys. JETP **38**, 494 (1974)].

The H$_2^+$ Ion Can Dissociate in a Strong Magnetic Field[1]

A. V. Turbiner[1,4], J.-C. López[2], and A. Flores-Riveros[3]

[1] *Laboratoire de Physique Theorique, Universite Paris-Sud, Orsay F-91405, France,*
and Instituto de Ciencias Nucleares, UNAM, México

[2] *Instituto de Ciencias Nucleares, UNAM, Apartado Postal 70-543, 04510, México D.F., México*
e-mail: vieyra@xochitl.nuclecu.unam.mx

[3] *Instituto de Fisica, Benemérita Universidad Autónoma de Puebla, Apartado Postal J-48, Puebla Pue. 72570, México*
e-mail: flores@sirio.ifuap.buap.mx

[4] *On leave from the Institute of Theoretical and Experimental Physics, Moscow, 117259 Russia*
e-mail: turbiner@xochitl.nuclecu.unam.mx

Received October 17, 2000

In the framework of a variational method, the molecular ion H$_2^+$ in a magnetic field is studied. The optimal form of the vector potential corresponding to a given magnetic field (gauge fixing) is chosen variationally. It is shown that for any magnetic field strength, as well as for any orientation of the molecular axis, the system (*ppe*) possesses a minimum in the potential energy. The stable configuration always corresponds to the elongation along the magnetic line. However, for magnetic fields $B \gtrsim 2 \times 10^{11}$ G and some orientations, the H$_2^+$ ion becomes unstable, decaying into the H atom +*p*. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 31.15.Fm; 31.15.Pf; 32.60.+i

The H$_2^+$ ion is simplest one-electron molecular system, which is more stable than the hydrogen atom. It appears to be one of the most studied problems in non-relativistic quantum mechanics. In particular, the wealth of physical phenomena displayed by this system in a magnetic field becomes of great importance in astrophysics, solid state, and plasma physics. For instance, as the magnetic field grows, the system becomes more and more strongly bound and compact. Such a behavior naturally leads to the guess that, in spite of the huge temperatures at neutron star surfaces, their atmosphere can still contain molecular objects [1, 2]. On the other hand, a shrinking of the equilibrium distance between protons with the growth of the magnetic field drastically increases the probability of nuclear fusion [3]. It is quite surprising that such a shrinking is also accompanied by a change from ionic to covalent character at ~5×10^{11} G [4] (see also [5]). The goal of this paper is twofold. Firstly, to show that the system (*ppe*) always has a minimum and, correspondingly, the molecular ion H$_2^+$ can exist at least in magnetic fields $\lesssim 4.414 \times 10^{13}$ G. Secondly, to demonstrate that for $B \gtrsim 2 \times 10^{11}$ G and for some orientations of the molecular axis the ion becomes unstable, dissociating into the H atom +*p*. The variational method is used to study this problem.

The Hamiltonian describing the H$_2^+$ molecular ion placed in a uniform constant magnetic field directed along the *z* axis, $\mathcal{B} = (0, 0, B)$ (see, for example, [6]) is given by

$$\mathcal{H} = \hat{p}^2 + \frac{2}{R} - \frac{2}{r_1} - \frac{2}{r_2} + (\hat{p}\mathcal{A}) + \mathcal{A}^2, \qquad (1)$$

(see Fig. 1), where $\hat{p} = -i\nabla$ is the momentum and $\mathcal{A}$ is a vector potential corresponding to the magnetic field $\mathcal{B}$.

The vector potential for a given magnetic field is defined ambiguously to a gauge factor. Thus, the Hamiltonian is gauge-dependent, but the observables are not. Since we are going to use an approximate method for solving Eq. (1), our energies can be gauge-dependent (only the exact ones would be gauge-independent). Hence, one can choose the form of the vector potential in an optimal way. Let us consider a certain one-parameter family of vector potentials corresponding to the constant magnetic field B (see, for example, [4])

$$\mathcal{A} = B(-(1 - \xi)y, \xi x, 0), \qquad (2)$$

where ξ is the parameter to be fixed in a certain optimal way. If $\xi = 1/2$, we get a gauge called symmetric or circular, while $\xi = 0$ corresponds to an asymmetric gauge

---

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

(see [6]). By substituting Eq. (2) into Eq. (1), we arrive at the Hamiltonian

$$\mathcal{H} = -\nabla^2 + \frac{2}{R} - \frac{2}{r_1} - \frac{2}{r_2} + iB[-(1-\xi)y\partial_x + \xi x\partial_y]$$
$$+ B^2[\xi^2 x^2 + (1-\xi)^2 y^2]. \tag{3}$$

The idea of choosing an optimal gauge has widely been exploited in quantum field theoretical considerations. It has also been discussed in connection with the problem at hand (see, for instance, [7] and references therein). Perhaps, the first constructive (and remarkable) attempt to apply this idea was made by Larsen [4]. In his study of the ground state, it was explicitly shown that the gauge dependence of energy can be quite significant and even an oversimplified optimization procedure improves the numerical results.

Our aim is to study the ground state for the Hamiltonian (3). It is not difficult to see that there exists a certain gauge for which Hamiltonian (3) has a real ground-state eigenfunction. This gauge will be sought below, and, correspondingly, we will deal with *real* trial functions in our variational calculations. In this case, one can prove that the expectation value for the term ~B in Eq. (3) vanishes when it is taken over any real normalizable function. So, without loss of generality we can omit this term in the Hamiltonian. Finally, the recipe of our variational study can be formulated as follows: *Construct an adequate variational real trial function [8] that reproduces the original potential near Coulomb singularities and at large distances, where ξ should be included as a parameter. Perform a minimization of the energy functional by treating the trial function's free parameters and ξ on the same footing*. In particular, such an approach enables one to eventually find the *optimal* form of the Hamiltonian. The above recipe was successfully applied in the previous study of H$_2^+$ in magnetic field [5] and allowed one to predict the existence of the exotic ion H$_3^{++}$ at $B \gtrsim 10^{11}$ G [9].

One of the simplest trial functions satisfying the above-mentioned criterion is

$$\Psi_1 = \exp\{-\alpha_1(r_1 + r_2)$$
$$- B[\beta_{1x}\xi x^2 + \beta_{1y}(1-\xi)y^2]\}, \tag{4}$$

(cf. [5]), where α_1 and $\beta_{1x, 1y}$ are the variational parameters. We assume that $\xi \in [0, 1]$, which is a restriction that will later be justified. Actually, this is a Heitler–London function multiplied by the lowest Landau orbital associated with gauge (2). This function presumably describes internuclear distances near the equilibrium and a covalent character. Another possible trial function is

$$\Psi_2 = (\exp(-\alpha_2 r_1) + \exp(-\alpha_2 r_2))$$
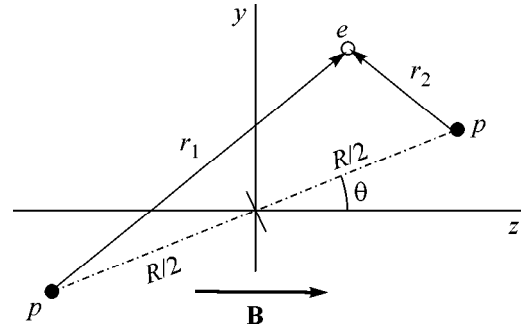$$\times \exp\{-B[\beta_{2x}\xi x^2 + \beta_{2y}(1-\xi)y^2]\}, \tag{5}$$

Fig. 1. Geometrical setting for the H$_2^+$ ion in a magnetic field directed along the z axis.
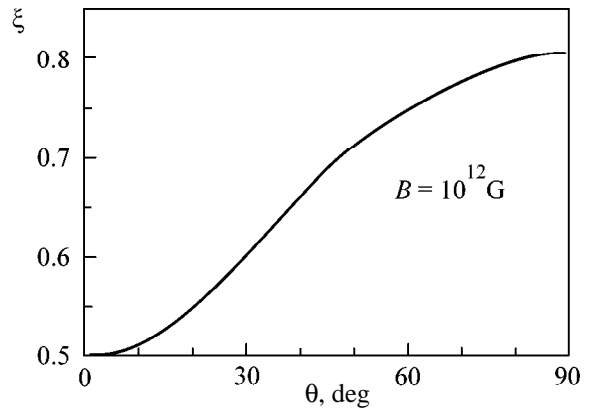

Fig. 2. Gauge parameter as a function of the inclination angle for H$_2^+$. The magnetic field $B = 10^{12}$ G was taken as an example. This dependence is found in the present study.

(cf. [5]), where α_2 and $\beta_{2x, 2y}$ are the variational parameters. This is a Hund–Mulliken function multiplied by the lowest Landau orbital. One can assume that for a sufficiently large internuclear distance R, this function dominates, thus giving an essential contribution in this regime. Hence, it describes an interaction of a hydrogen atom and a proton (charged center) and can also describe a possible decay mode of H$_2^+$ into them. There are two natural ways—linear and nonlinear—to incorporate the behavior of the system both near equilibrium and at large distances in a single trial function. The nonlinear interpolation is of the form

$$\Psi_{3-1} = (\exp\{-\alpha_3 r_1 - \alpha_4 r_2\} + \exp\{-\alpha_3 r_2 - \alpha_4 r_1\})$$
$$\times \exp\{-B[\beta_{3x}\xi x^2 + \beta_{3y}(1-\xi)y^2]\} \tag{6}$$

(cf. [5]), where $\alpha_{3, 4}$ and $\beta_{3x, 3y}$ are the variational parameters. This is a Guillemin–Zener function multiplied by the lowest Landau orbital. The linear superposition is given by
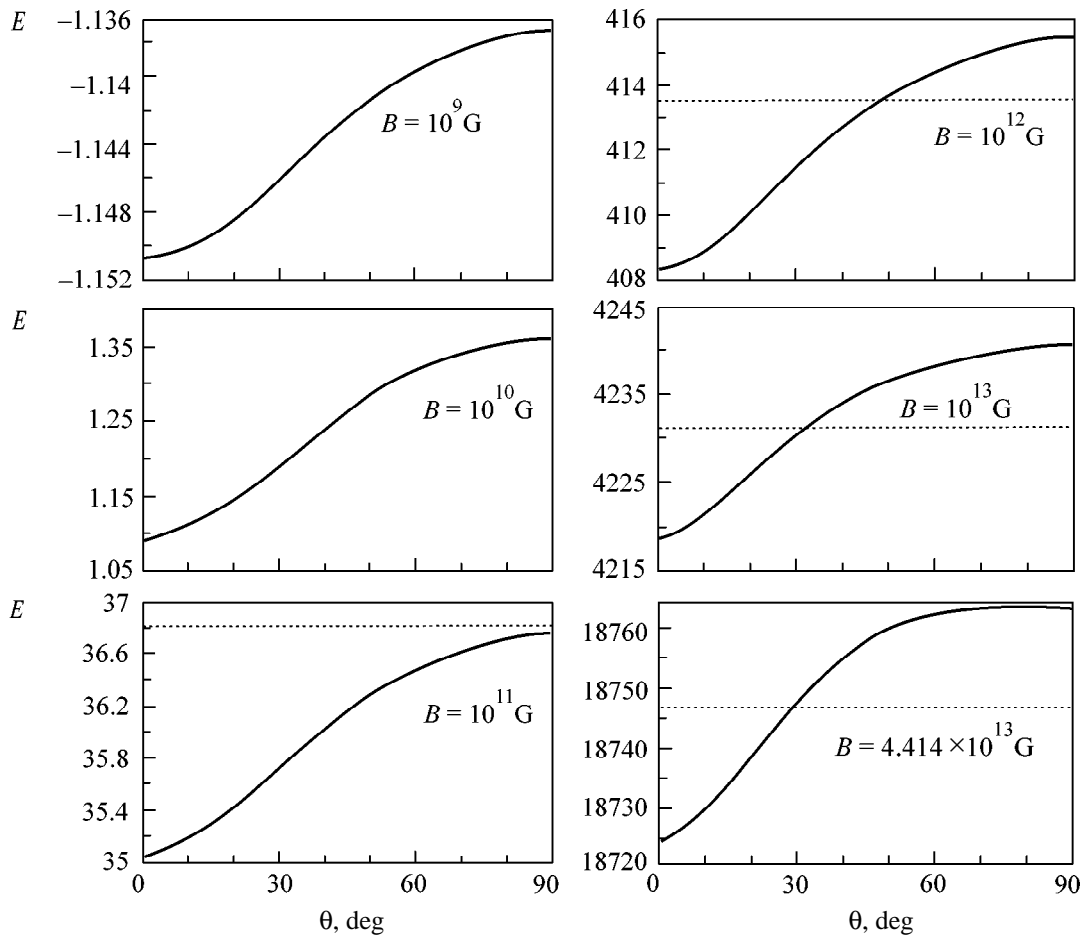
$$\Psi_{3-2} = A_1\Psi_1 + A_2\Psi_2, \tag{7}$$

Fig. 3. Total energy of the H_2^+ ion as a function of the inclination angle. The horizontal lines refer to the H ground-state energy taken from Lai *et al.* [11].

where one of the parameters $A_{1, 2}$ is kept fixed. The final form of the trial function is a linear superposition of functions (6) and (7)

$$\Psi_{trial} = A_1\Psi_1 + A_2\Psi_2 + A_{3-1}\Psi_{3-1}, \qquad (8)$$

where only two of the three A parameters are variationally treated. The total number of variational parameters in Eq. (8) is fourteen when ξ is included.

It is easy to prove the general statement that if a system possesses axial rotational symmetry (in our case, it appears if the molecular axis coincides with the magnetic line, $\theta = 0°$; see Fig. 1), the optimal gauge corresponds to $\xi = 1/2$ (symmetric or circular gauge). It is precisely this gauge which was used in most of previously performed H_2^+-studies. However, this is not the case if $\theta \neq 0°$. As an example, one can see in Fig. 2 the behavior of ξ as a function of θ at $B = 10^{12}$ G. It is typical behavior for all studied magnetic fields, both weak and strong, up to the nonrelativistic limit $B = 4.414 \times 10^{13}$ G. It justifies our above assumption with regard to the domain of gauge parameter $\xi \in [0, 1]$.

We carried out extensive studies of the ground state of H_2^+ for magnetic fields $B = 0–4.414 \times 10^{13}$ G and orientations ranging from $0°$ (parallel configuration) to $90°$ (perpendicular configuration). They turned out to be more accurate than any available results obtained to date, except for a domain of small magnetic fields for the perpendicular configuration, where the results of Wille [10] appear to be slightly better.[2] The detailed numerical analysis and comparison with available calculations will be published elsewhere.

As previously obtained by other authors [3, 10, 4], we quantitatively confirm the natural expectation that the parallel configuration is the most stable for all magnetic fields $B \leq 4.414 \times 10^{13}$ G, where nonrelativistic considerations are valid. The total energy of the molecular ion H_2^+ as a function of angle θ for different magnetic fields is shown in Fig. 3. For any magnetic field in the region

---

[2] These results were obtained using a basis-set expansion including up to approximately 500 terms, depending on the value of magnetic field.
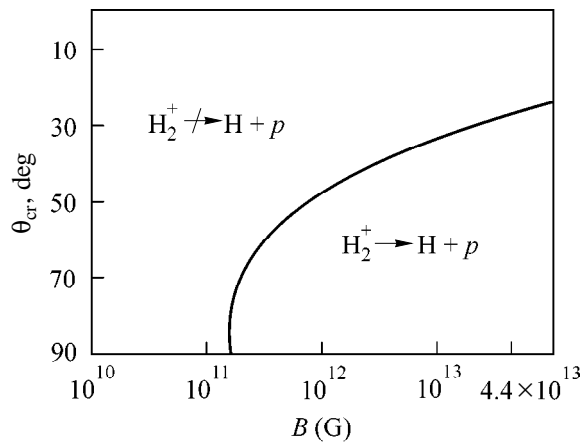
Fig. 4. Critical angle for dissociation of H$_2^+$.

$B = 0$–4.414×10^{13} G and any orientation, a well-pronounced minimum in the total energy is attained at finite internuclear distances. This is in contradiction with a statement by Khersonsky [3] about the nonexistence of a minimum for some values of the magnetic field and orientation. Perhaps, it should be emphasized that the variational study in that article was carried out using a trial function which is almost coincident with that in Eq. (5). We can thus presume that the above statement is an artifact arising from insufficient accuracy of calculations. The horizontal line in Fig. 3 presents the hydrogen atom total energy in the magnetic field (see [11]). For magnetic fields $B \gtrsim 1.8 \times 10^{11}$ G, the total energy of the atom becomes lower than that of H$_2^+$ for angles larger than some critical angle, θ_{cr}. For $\theta = 90°$, a similar statement was made by Larsen [4]. It points to the possible dissociation channel H$_2^+ \longrightarrow$ H atom $+ p$. The dependence of the critical angle θ_{cr} on the magnetic field is shown in Fig. 4. It is quite striking that dissociation occurs for a wider and wider range of orientations as the magnetic field grows, reaching $25° \lesssim \theta \leq 90°$ for $B = 4.414 \times 10^{13}$ G.

REFERENCES

1. B. B. Kadomtsev and V. S. Kudryavtsev, Pis'ma Zh. Éksp. Teor. Fiz. **13**, 15 (1971) [JETP Lett. **13**, 9 (1971)]; Pis'ma Zh. Éksp. Teor. Fiz. **13**, 61 (1971) [JETP Lett. **13**, 42 (1971)]; Zh. Éksp. Teor. Fiz. **62**, 144 (1972) [Sov. Phys. JETP **35**, 76 (1972)].

2. M. Ruderman, Phys. Rev. Lett. **27**, 1306 (1971); in *Proceedings of 53rd IAU Symposium on Physics of Dense Matter*, Ed. by C. J. Hansen (Reidel, Dordrecht, 1974), p. 117.

3. V. Khersonsky, Astrophys. Space Sci. **87**, 61 (1982); **98**, 255 (1984).

4. D. Larsen, Phys. Rev. A **25**, 1295 (1982).

5. J.-C. López, P. O. Hess, and A. Turbiner, Phys. Rev. A **56**, 4496 (1997); astro-ph/9707050.

6. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon, New York, 1977, 3rd ed.).

7. P. Schmelcher, L. S. Cederbaum, and U. Kappes, in *Conceptual Trends in Quantum Chemistry* (Kluwer, Dordrecht, 1994).

8. A. V. Turbiner, Zh. Éksp. Teor. Fiz. **79**, 1719 (1980) [Sov. Phys. JETP **52**, 868 (1980)]; Usp. Fiz. Nauk **144**, 35 (1984) [Sov. Phys. Usp. **27**, 668 (1984)]; Yad. Fiz. **46**, 204 (1987) [Sov. J. Nucl. Phys. **46**, 125 (1987)]; Doctoral Dissertation (Inst. Teor. Éksp. Fiz., Moscow, 1989).

9. A. Turbiner, J.-C. Lopez, and U. H. Solis, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 800 (1999) [JETP Lett. **69**, 844 (1999)]; astro-ph/9809298.

10. U. Wille, Phys. Rev. A **38**, 3210 (1988).

11. D. Lai, E. Salpeter, and S. L. Shapiro, Phys. Rev. A **45**, 4832 (1992).

Chaotic Motion of Atom in the Coherent Field of a Standing Light Wave

S. V. Prants and L. E. Kon'kov

Il'ichev Pacific Oceanological Institute, Far East Division, Russian Academy of Sciences,
ul. Baltiĭskaya 43, Vladivostok, 690041 Russia
Received October 2, 2000; in final form, December 26, 2000

A new effect of chaotic motion of the center of mass of a cold atom in the coherent field of a standing light wave in a high-finesse Fabry–Pérot cavity is theoretically predicted and numerically implemented in the absence of any random fluctuations due to spontaneous emission. Numerical experiments demonstrate that the Hamiltonian chaos arises near resonance in the range of parameters characteristic of the strong coupling regime that was implemented in recent experiments. The effect is of interest in studying the quantum–classical correspondence and quantum chaos in atomic optics. © *2001 MAIK "Nauka/Interperiodica"*.

PACS numbers: 32.80.Lg; 42.50.Vk; 42.65.Sf; 05.45.Mt

The interaction of a free atom with a resonance electromagnetic field not only changes the internal (electronic) degrees of freedom of an atom but simultaneously changes its external (translational) degrees of freedom. The ideas of laser action on the translational degrees of freedom [1–4] have led to the development of powerful methods of laser-induced atom cooling and trapping [5–7]. Under ordinary conditions, the light intensity is high enough so that it can be considered as an inexhaustible energy and momentum reservoir for an atom whose presence has no effect on the field state. In a high-finesse microcavity, the interaction of an atom with the field is strong to the extent that a multiple excitation exchange between them becomes possible. The strong coupling regime for the internal ($\Omega_0 \sqrt{n} > \Gamma_{a,f}^{-1}$, where Ω_0 is the single-photon Rabi frequency, n is the mean number of photons in the mode, and $\Gamma_{a,f}$ are the relaxation rates of atomic dipole and photons, respectively) and external ($\hbar\Omega_0 \sqrt{n} > m v_a^2$, where v_a is the atomic velocity) degrees of freedom has recently been realized in [8, 9].

A fundamental model of interaction of the atomic internal and external degrees of freedom with the field of a standing light wave is expressed by the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}\hbar\omega_a\hat{\sigma}_z + \hbar\omega_f\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)$$
$$+ \hbar\Omega_0(\hat{a}^\dagger\hat{\sigma}_- + \hat{a}\hat{\sigma}_+)\sin(k\hat{x} + \varphi) \quad (1)$$

that describes, in the rotating-wave approximation, the coherent dynamics of a two-level atom placed in an ideal one-dimensional Fabry–Pérot-type cavity. In this communication, it is shown that: (i) within the frame-

work of the fundamental model, a new effect of chaotic motion of the center of mass of an atom in the coherent potential field of a standing light wave is possible in the absence of any random fluctuations; (ii) the Hamiltonian dynamic chaos arises near the resonance in the range of parameters characteristic of strong coupling regime [8, 9]; and (iii) this effect leads to partial destruction of a regular structure formed by the potential wells (traps) for slow atoms and to the disappearance of a regular spatial modulation of Rabi oscillations for fast atoms. It should be emphasized that the effect of chaotic motion of the center of mass of an atom arises in model (1) describing the dynamic interaction of the atomic internal and external degrees of freedom with the field degree of freedom. So far, the chaos in the interaction of the elementary quantum systems with radiation has been studied using models disregarding a change in the momentum of the system in the course of interaction with field (see pioneer work [10] and other papers, e.g., [11, 12] and literature cited therein). In [13, 14], the dynamic chaos was numerically revealed and studied, with allowance made for the spatial structure of standing wave, but again without regard for a change in the atomic momentum, within the framework of a model of parametric interaction between the moving atoms and the field (i.e., in the Raman–Nath approximation).

Hamiltonian (1) describes the interaction between three subsystems: translational, characterized by the atomic coordinate $\hat{x}$ and momentum $\hat{p}$ operators; electronic, characterized by the Pauli matrices $\hat{\sigma}_\pm = \frac{1}{2}(\hat{\sigma}_x \pm i\hat{\sigma}_y)$ and $\hat{\sigma}_z$; and the field subsystem, characterized by the photon annihilation $\hat{a}$ and creation $\hat{a}^\dagger$ operators. The semiclassical coherent dynamics of

interaction between these systems is described by the Heisenberg equations

$$\dot{\xi} = \alpha\rho, \quad \dot{\rho} = -u\sin\xi, \quad \dot{u} = \delta v,$$
$$\dot{v} = -\delta u + 2Nz\cos\xi, \quad \dot{z} = -2v\cos\xi \quad (2)$$

that are derived from Hamiltonian (1) with initial phase $\varphi = -\pi/2$ for the following dimensionless variables:

$$\xi = k\langle\hat{x}\rangle, \quad \rho = \frac{1}{\hbar k}\langle\hat{p}\rangle,$$
$$u = \langle\hat{a}^\dagger\hat{\sigma}_- + \hat{a}\hat{\sigma}_+\rangle, \quad v = i\langle\hat{a}^\dagger\hat{\sigma}_- - \hat{a}\hat{\sigma}_+\rangle, \quad (3)$$
$$z = \langle\hat{\sigma}_z\rangle,$$

which are averages of the corresponding operators in the quantum state $|\psi\rangle$ of the total system. It is assumed to be factorized at the initial time $\tau_0 = \Omega_0 t_0$ in a product of translational $|\psi_0\rangle_{tr}$, electronic $|\psi_0\rangle_e$, and field $|\psi_0\rangle_f$ vector states. The normalized resonance detuning $\delta = (\omega_f - \omega_a)/\Omega_0$, the normalized recoil frequency $\alpha = \hbar k^2/m\Omega_0 \equiv 2\omega_R/\Omega_0$, and the number of excitations $N = n + \frac{1}{2}(z+1)$, which, as known, is conserved in the rotating-wave approximation, are the driving parameters of the closed system of nonlinear Eqs. (2). System of Eqs. (2) provides the conservation of the total energy and the electron–field interaction energy,

$$W = \frac{\alpha}{2}\rho^2 - u\cos\xi - \frac{\delta}{2}z, \quad R = u^2 + v^2 + Nz^2. \quad (4)$$

It can be interpreted as a Hamiltonian system with the coupled translational (the ξ and ρ variables) and electron–field (the u and v variables) degrees of freedom. At the initial time, the u and v quantities are simple combinations $u_0 = (e_0 x_0 - p_0 y_0)/2$ and $v_0 = (e_0 y_0 + p_0 x_0)/2$ of the field variables $e_0 \equiv \langle\hat{a}^+ + \hat{a}\rangle_f$ and $p_0 \equiv i\langle\hat{a}^+ - \hat{a}\rangle_f$ and the atomic dipole-moment variables $x_0 \equiv \langle\hat{\sigma}_x\rangle_e$ and $y_0 \equiv \langle\hat{\sigma}_y\rangle_e$.

If the mode frequency ω_f coincides with the atomic transition frequency ω_a, i.e., if $\delta = 0$, then it follows from the third equation in Eqs. (2) that there is an additional integral of motion $u = u_0$. In this case, the atomic center-of-mass motion in the potential periodic field $U = -u_0\cos\xi$ produced by the standing wave obeys the simple equation of a free nonlinear oscillator $\ddot{\xi} + \alpha u_0\sin\xi = 0$ and is independent of the internal atomic state. Depending on the initial conditions, the atom is either trapped by the standing wave, so that its center of mass executes periodic oscillations, or crosses the standing wave in such a manner that its momentum is periodically modulated. The atom is trapped by the field if its initial momentum satisfies the condition $\rho_0 \leq \rho_{cr} = 2\sqrt{u_0/\alpha}$. The critical atomic velocity and the nor-

malized frequency $\sqrt{\alpha u_0}$ of atomic small-amplitude oscillations in the potential well depend on the initial field and atomic dipole components. It follows from these relationships that, in the exact resonance, the atom moves uniformly and rectilinearly, provided that the field and/or atom are initially prepared in such a manner that $u_0 = 0$, i.e., if $e_0 x_0 = p_0 y_0$. In particular, the completely inverted atom ($z_0 = 1$ and $x_0 = y_0 = 0$) will continue the uniform and rectilinear motion irrespective of the initial standing-wave state $|\psi_0\rangle_f$. The same will occur with an atom entering the cavity in which the field is in the Fock's state $|n\rangle_f$ with a specified number of photons, regardless of the initial state $|\psi_0\rangle_e$ of the atom. Clearly, these considerations are valid only in the absence of dissipation. The decay of the chosen mode and the spontaneous emission to other modes are ignored in the strong-coupling approximation.

On the other hand, the evolution of the atomic internal energy and the evolution of the average number of photons in the mode depend on the translational degrees of freedom, because the coupling strength between the atom and field depends on the atomic position in the periodic potential of the standing wave. The equation of motion for the population inversion can be obtained from the two last equations in Eqs. (2) with the use of Eqs. (4),

$$\dot{z} = \pm 2\sqrt{R - Nz^2}\cos\xi(\tau). \quad (5)$$

The solution of this equation is straightforward:

$$z(\tau) = \pm\sqrt{\frac{R}{N}}\sin\left(2\sqrt{N}\int_0^\tau\cos[\xi(\tau)]d\tau + \text{const}\right). \quad (6)$$

Therefore, the atomic internal energy (and the average number of photons in the mode) undergo periodically modulated oscillations. In the limit of large detunings, $\delta \gg \sqrt{N}$, the population inversion can be adiabatically eliminated from Eqs. (2). As a result, the number of degrees of freedom in the system of Eqs. (2) is effectively reduced to unity and it becomes integrable with the periodic solutions. In actuality, we are dealing with two coupled oscillators, one of which (electron–field) is linear, with the eigenfrequency δ being much higher than the frequency $\sqrt{\alpha u_0}$ of the small-amplitude translational oscillations.

In the limiting cases of large and small detunings, the system executes periodic motions. A simple qualitative analysis of equations of motion (2) yields the following quantitative criteria for the limiting cases: $\delta > \max\{1, 2\sqrt{N}, \alpha\rho_0/\sqrt{N}\}$ and $\delta < \min\{1, 2\sqrt{N}, \alpha\rho_0/\sqrt{N}\}$. The dynamical complexity in the interaction between the internal and external degrees of freedom of a two-level atom in the field of a standing light wave arises in the near-resonance region. Before proceeding to the numerical experiments, let us estimate the range
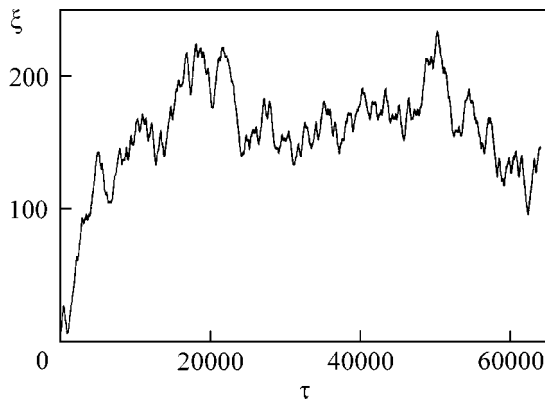
Fig. 1. Typical chaotic trajectory of an atom in the coherent field of a standing light wave.
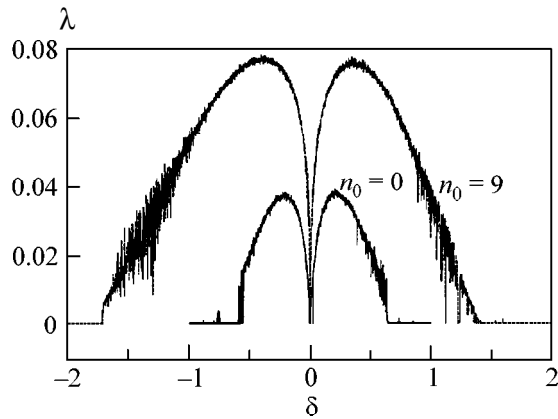

Fig. 2. The maximum Lyapunov exponent λ, in units of inverse dimensionless time, as a function of the normalized resonance detuning δ for two values of the average number of photons n_0 in the mode.

of driving parameters of our system in the strong-coupling regime. In experiments [8, 9], atoms (precooled in a magneto-optical trap down to the energy E_k in the millikelvin range) enter a high-finesse Fabry–Pérot microcavity and move in a radial plane (where the standing-wave mode is Gaussian-shaped) across the cavity axis x with a $\cos kx$ mode shape. At small velocities, the atoms are trapped by the Gaussian potential well. The real-time detection of the atomic oscillations in the well is accomplished by monitoring the intensity variations of a probe laser beam directed along the x axis. Since the probability of laser-induced transition between the states of the dressed atom depends on the atomic coordinate **r**, the intensity of light passed through the cavity with an atom inside it directly depends on the coupling strength $\Omega(\mathbf{r})$ between the atom and the field. The measured intensity oscillations can be used to reproduce the atomic trajectory in the radial plane.

The authors of [8] used cesium atoms with mass $m_a \simeq 10^{-22}$ g, kinetic energy $E_k \simeq 0.46$ mK, a working

transition frequency $\omega_a \simeq 2\pi \times 3.5 \times 10^8$ MHz, detuning $|\omega_f - \omega_a| \simeq 2\pi \times 50$ MHz, and decay rate of $\Gamma_a \simeq 2\pi \times 2.6$ MHz in a Fabry–Pérot cavity with a length of 10.9 μm, finesse $Q \simeq 4.8 \times 10^5$, field decay rate $\Gamma_f \simeq 2\pi \times 14.2$ MHz, and a standing-wave potential well depth of $\simeq 2.3$ mK. Since the single-photon Rabi frequency $\Omega_0 \simeq 2\pi \times 110$ MHz in the standing-wave antinode is larger than $\Gamma_{a,f}$ and $\hbar\Omega_0 \simeq 5.3$ mK is larger than E_k, the conditions for strong coupling are fulfilled for both internal and external atomic degrees of freedom. These quantities can be used to calculate the dimensionless driving parameters of our system [Eqs. (2)]: the recoil frequency $\alpha \simeq 4 \times 10^{-6}$ and the detuning $|\delta| \simeq 0.4$. The experimental parameters used in [9] for the ^{85}Rb atom give $\alpha \simeq 4.4 \times 10^{-4}$ and $|\delta| \simeq 1$, and $\alpha \simeq 10^{-3}$ and $|\delta| \simeq 0.5$ is obtained for the light He atom. In the semiclassical approximation, the atom is a material point with coordinate $\langle \hat{x} \rangle$ and momentum $\langle \hat{p} \rangle$ driven by the $-\hbar k\Omega_0 \langle \hat{u} \cos k\hat{x} \rangle$ force. This approximation is valid if the recoil energy $\hbar\omega_R$ is much lower than the other characteristic energies, i.e., the kinetic energy $mv_a^2/2$ and electron–field energy $\hbar\Omega_0$. It follows that the criterion for the validity of semiclassical approximation can be formulated as $\alpha \ll 1$ and $v_R = \hbar k/m \ll v_a$.

If the atom is completely inverted ($z_0 = 1$) at the initial time, then $u_0 = v_0 = 0$, regardless of the type of quantum state of a field in the cavity. At the exact resonance ($\delta = 0$), this atom crosses the standing light wave without being affected by the field. The situation drastically changes for the atom slightly detuned to both the red ($\delta < 0$) and blue ($\delta > 0$) regions. A typical trajectory of the He atom with initial momentum $\rho_0 = 50$ and $\delta = 0.5$ in the field of a standing wave with initial average number of photons $n_0 = 9$ is depicted in Fig. 1. When moving across the cavity, the atom oscillates in the wells and jumps at random from one well to the other. The fact that this motion is chaotic in the sense of exponential sensitivity to the changes in initial conditions is proved by evaluating the maximum Lyapunov exponent λ for the system of Eqs. (2) in units of inverse dimensionless time. Figure 2 shows the dependence of λ on the detuning δ for different initial values n_0. As expected, the chaos disappears ($\lambda = 0$) near the resonance and away from it. Note that the atom can chaotically move at $\delta \neq 0$ even if the field in the cavity is initially in the vacuum state. Figure 3 is the so-called topographical λ map [13, 14] constructed in the $v_a - \omega_R/2\pi$ coordinates for fixed $\delta = 0.5$ and $N = 10$. In Fig. 3, different regions of the λ values are differently shaded for the initial velocities v_a lying in the range 1–500 cm/s and the recoil frequencies $\omega_R/2\pi$ lying in the range 1–1000 kHz. This map is used to determine the values of indicated quantities (and, hence, the type of atom) for which one can expect either the chaotic or a regular motion of an atom in cavity.
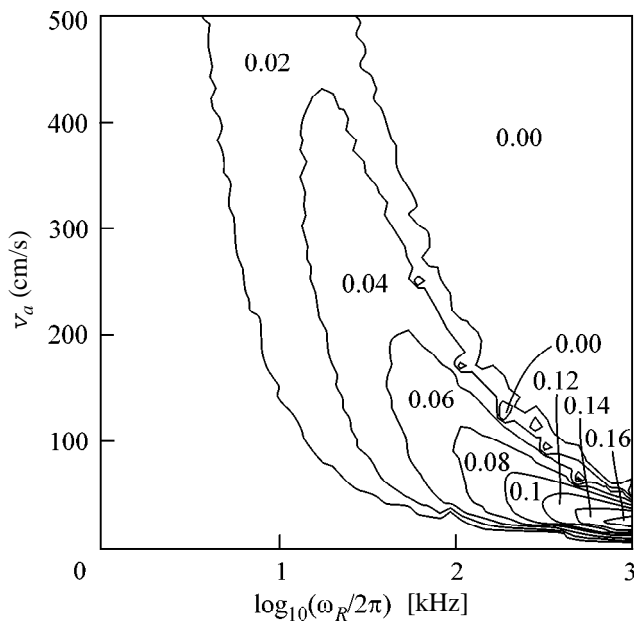
Fig. 3. The topographical map of λ values as a function of the initial atomic velocity v_a (cm/s) and the recoil frequency $\omega_R/2\pi$ (kHz).

In summary, we have predicted, both theoretically and numerically, the effect of chaotic motion of an atom in a high-finesse Fabry–Pérot cavity in the range of parameters corresponding to the strong-coupling regime. The semiclassical Hamiltonian chaos arises as a result of the coherent energy exchange between the translational and electron–field degrees of freedom of a nonlinear atom–field–cavity system in the absence of any random fluctuations due to spontaneous emission or to the other coherence-breaking factors. The realization of the strong-coupling regime for single atoms and photons in experiments [8, 9] allows one to believe that special experiments on studying the quantum–classical correspondence and the quantum chaos in atomic optics would be quite promising. The studies of quantum dynamics for an atomic wave packet in the quantized field and, in particular, of the effect of dynamical localization in the range of parameters where we have revealed semiclassical chaos seem to be of particular interest.

Note in conclusion that the Hamiltonian dynamics studied in this work for an atom in the spatially periodic potential can be considered adequate in the extremely strong-coupling regime ($\Omega_0 \sqrt{n} \gg \Gamma_{a,f}$). Rapid advance in experimental technique allows us to believe that this regime will be realized in near future for single atoms and photons in the optical range. The dissipation and spontaneous emission can be taken into account within the framework of semiclassical approach. It is conceivable that this will give rise to a dissipative chaos, which deserves special investigation.

REFERENCES

1. G. A. Askar'yan, Zh. Éksp. Teor. Fiz. **42**, 1567 (1962) [Sov. Phys. JETP **15**, 1088 (1962)].
2. V. S. Letokhov, Pis'ma Zh. Éksp. Teor. Fiz. **7**, 348 (1968) [JETP Lett. **7**, 272 (1968)].
3. A. Ashkin, Phys. Rev. Lett. **25**, 1081 (1970).
4. T. W. Hansch and A. L. Shawlow, Opt. Commun. **13**, 68 (1975).
5. S. Chu, Rev. Mod. Phys. **70**, 685 (1998).
6. C. Cohen-Tannoudji, Rev. Mod. Phys. **70**, 707 (1998).
7. W. D. Phillips, Rev. Mod. Phys. **70**, 721 (1998).
8. J. Yen, D. W. Vernooy, and H. J. Kimble, Phys. Rev. Lett. **83**, 4987 (1999); C. J. Hood, T. W. Lynn, A. C. Doherty, *et al.*, Science **287**, 1447 (2000).
9. P. Münstermann, T. Fischer, P. Maunz, *et al.*, Phys. Rev. Lett. **82**, 3791 (1999).
10. P. I. Belobrov, G. M. Zaslavskiĭ, and G. Kh. Tartakovskiĭ, Zh. Éksp. Teor. Fiz. **71**, 1799 (1976) [Sov. Phys. JETP **44**, 945 (1976)].
11. J. R. Ackerhalt, P. W. Milonni, and M. L. Shih, Phys. Rep. **128**, 205 (1985).
12. A. R. Kolovsky, S. Miyazaki, and R. Graham, Phys. Rev. E **49**, 70 (1994).
13. S. V. Prants and L. E. Kon'kov, Phys. Lett. A **225**, 33 (1997).
14. S. V. Prants and L. E. Kon'kov, Phys. Rev. E **61**, 3632 (2000).

Translated by V. Sakun

X-ray Microscopy with Asymmetrical Reflection
from a Single Crystal

A. V. Andreev[1], V. E. Asadchikov[2], B. V. Mchedlishvili[2], Yu. V. Ponomarev[1],
A. A. Postnov[2], R. A. Senin[1], and T. V. Tsyganova[2]

[1] *Moscow State University, Vorob'evy gory, Moscow, 119899 Russia*
[2] *Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninskiĭ pr. 59, Moscow, 117333 Russia*
Received January 25, 2001

The theoretical and experimental possibilities of obtaining enlarged images in the hard X-ray range with the use of asymmetrical crystals (monochromators) are considered. © *2001 MAIK "Nauka/Interperiodica".*

PACS numbers: 41.50.+h; 07.85.Tt

At present, a considerable interest is being shown, both in Russia and abroad, in works devoted to obtaining images of internal details of nontransparent objects by the methods of X-ray microscopy [1]. Depending on the chosen wavelength, the use of X-ray radiation allows one to vary both the radiation penetration depth into a substance and the contrast of various chemical compounds and biological tissues relative to each other. Unfortunately, one cannot as yet realize the huge resolving potentialities of X-ray optics because of its poor development (the theoretically highest attainable resolution is comparable to the wavelength), especially for energies higher than 1 keV. The potentialities of X-ray microscopy without the use of optics are limited by a resolution of several microns because of either the finite sizes of a source [2] (in projection microscopy) or a poor detector resolution [3] (in contact microscopy). Attempts at improving the resolution to fractions of a micron lead to large exposure times (tens of hours), thereby restricting the practical use of these microscopy schemes.

In this work, we use a range of hard X-rays ($\lambda = 0.05$–0.3 nm). In our opinion, this wavelength range is promising for studying various objects, because the absorption and refraction are comparatively weak for most materials, whereas the contrast is still noticeable. In addition, the use of such radiation does not necessitate the evacuation of an apparatus; i.e., the proposed scheme can be realized in the laboratory conditions with the use of an X-ray tube as a source.

In this wavelength range, Fresnel zone plates [4] for transmission [4] and reflection (Bragg–Fresnel optics [5]) have found the greatest application. The fabrication of such optics is a highly complicated technological problem, while the resolution, determined by the width of external zone, presently does not exceed 0.2 μm for the best lithographically fabricated zone plates [6]. At present, these zone plates are used only for synchrotron radiation sources. The corresponding

visual field is equal to hundreds of microns. In this paper, the reflection from an asymmetrical crystal (monochromator) is suggested for obtaining enlarged images [7]. Such a crystal may become an alternative optical element for laboratory X-ray microscopes, because it provides (as will be shown below) a rather high resolution and a large visual field. This paper is devoted to the estimates of theoretical resolution and the description of our experiments with an asymmetrical silicon crystal.

Theoretical estimates of resolution. Let us prove that the submicron resolution and a 30–50-fold enlargement can be obtained by using the above-mentioned radiation and the asymmetrical reflection from perfect single crystals. As known, the linear dimensions d_r of the reflected beam become different from the incident value d_0 upon the asymmetrical Bragg reflection. In geometrical optics, their ratio is given by the expression

$$d_r/d_0 = \sin\theta_r/\sin\theta_0, \tag{1}$$

where θ_0 is the angle of incidence and θ_r is the angle of reflection.

Let us demonstrate that the micron resolution can be obtained with the use of ordinary X-ray tubes if the crystal parameters and the wavelength of incident radiation are optimally chosen. Equation (1) can be used to estimate the maximum enlargement for a single reflection. Setting $\sin\theta_r \approx 1$ and taking into account that the smallest angle of incidence for a highly asymmetrical diffraction is determined by the angle of total external reflection $\theta_c = \sqrt{|\chi_0|} \approx 4 \times 10^{-3}$ rad, one finds that the maximum enlargement is on the order of 250. However, Eq. (1) is only valid under the condition that the beam widths are much larger than the extinction length. For an infinitely narrow incident beam, the characteristic sizes of the reflected beam are determined by the projection of the extinction length onto the entrance face of a crystal. It is conceivable that it is this fact which

reduces interest in the use of asymmetrical reflection in X-ray microscopy. Let us determine the limiting possibilities for spatial resolution in this scheme. In the diffraction by a perfect crystal, the extinction length of X-ray radiation is given by the expression $l_c = \lambda/|\chi_h|$. For the CuK$_\alpha$ radiation, this is on the order of 20 μm in silicon and 5–10 μm in germanium (χ_0 and χ_h are the Fourier components of dielectric susceptibility). For an angle of incidence of 10^{-2} rad and hole sizes of 0.2 μm in the object, its projection onto the crystal surface is 20 μm; i.e., it coincides with the extinction length in silicon and is 2–4 times larger than the extinction length in germanium. For instance, for a germanium (111) crystal, such angle gives a value of 40 for the asymmetry parameter, so that the size of the spot on photographic film will be 8 μm. With the photographic film, this is likely close to the limiting value, because the grain size in the film is 2–5 μm. However, the resolution of the scheme, in principle, can be brought to 1000 Å upon a twofold decrease in the angle of incidence. In Fig. 1, the curve for the total external reflection is compared with the curve for a diffraction reflection. It is seen in this figure that the intensity of the diffracted wave exceeds the intensity of the reflected wave at the angle of incidence $\theta_c \le \theta_0 \le 2\theta_c$, so that the reflection losses are insignificant. Let us estimate the diffraction limits of reflection. The angular divergence caused by the passage of a plane wave through the hole can be expressed by the formula $\Delta\theta = \lambda/d$. The corresponding linear expansion of a beam at distance l is $\Delta x = l\Delta\theta$. Consequently, the divergence of a plane wave with $\lambda \approx 0.1$ nm will be equal to $\Delta\theta \approx 10^{-4}$ after passing through a hole with diameter $d = 1$ μm, so that the holes spaced at $\Delta x = 1$ μm will merge at a distance of $l = 1$ cm.

The angular divergence of a polychromatic beam of the X-ray tube should not exceed the divergence due to the diffraction by the object, so that, if we want to obtain a resolution of 1 μm, it should be less than 10^{-4}. This limitation is not stringent in the single-reflection schemes and can easily be removed by choosing the appropriate distance from the tube to the sample under study and the appropriate tube focus size. However, the use of a scheme with two reflections in mutually perpendicular planes is more efficient for the two-dimensional enlargement. Due to the dispersion properties of this scheme, one can make the distance from the tube to the sample arbitrarily short, thereby appreciably enhancing the intensity of the reflected beam and reducing the exposure time.

Enlarged images experimentally obtained using an asymmetrical monochromator. The scheme of our experimental setup is shown in Fig. 2. An air-cooled point source of MoK$_\alpha$ radiation was used (the diameter of a focal spot was about 50 μm). The widths of slit diaphragms 2 were chosen to be 250 μm. The sample was fixed at the second diaphragm as close to the asymmetrical monochromator as possible. The monochromator was mounted on a goniometer and could be moved in
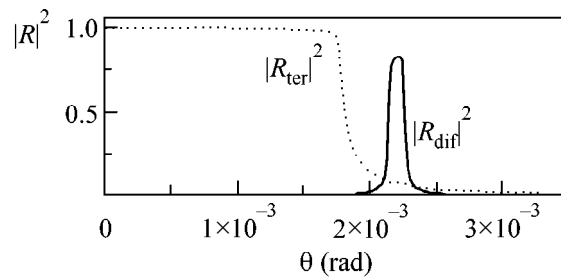
Fig. 1. The coefficients of diffraction ($|R_{\rm dif}|^2$) and total external ($|R_{\rm ter}|^2$) reflections as functions of glancing angle.
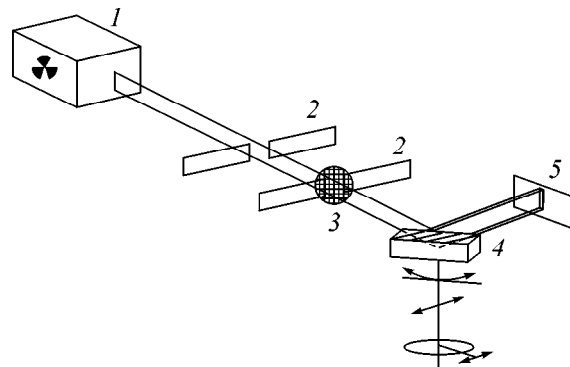
Fig. 2. Scheme of experimental setup. (*1*) Microfocal source, (*2*) collimating slits, (*3*) sample, (*4*) asymmetrical crystal on a goniometer, and (*5*) photographic film.

the direction perpendicular to the incident radiation and rotated about two perpendicular axes (one of them was aligned with the normal to the entrance face of crystal). To reduce the resolution losses caused by beam divergence, the object and the detector were be placed as close to the crystal as possible. A Si crystal with an asymmetrical (111) reflection was used as the enlarging optical element. The angle between the crystal surface cut and the reflecting crystallographic planes was 5.64°. According to Eq. (1), this provided a ~20-fold enlargement (in one direction, because only one monochromator was used). The crystal tuning to the reflection was carried out using a scintillation detector. The enlarged image of the object was placed on the photographic film UFShS (*Khimfotoproekt* Research Institute, Russia). Copper objective gauze for electron microscopy and a tantalum impact membrane were chosen as objects for investigation. The electron-microscopic images of the objects are shown in Fig. 3. One can see that the objective gauze for electron microscopy is a network of square holes with sides of ~65 μm and a spacing of ~100 μm. The impact track membranes represent a system of holes with diameters of 10–70 μm in a tantalum foil. These holes (tracks) may be differently inclined to the membrane surface (Fig. 3, tracks a, b).
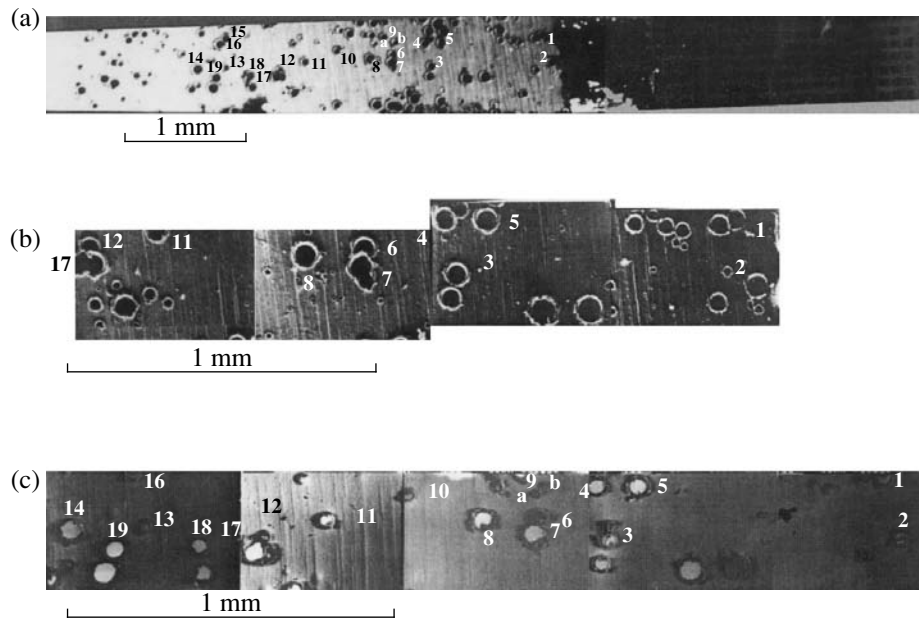
Fig. 3. Electron-microscopic images of the objects studied: (a) and (b) are images of the surfaces of track membranes, as obtained with different enlargements on the side of particle entrance, and (c) is the same on the exit side. The image of electron-microscopic gauze is seen at the right of panel (a).
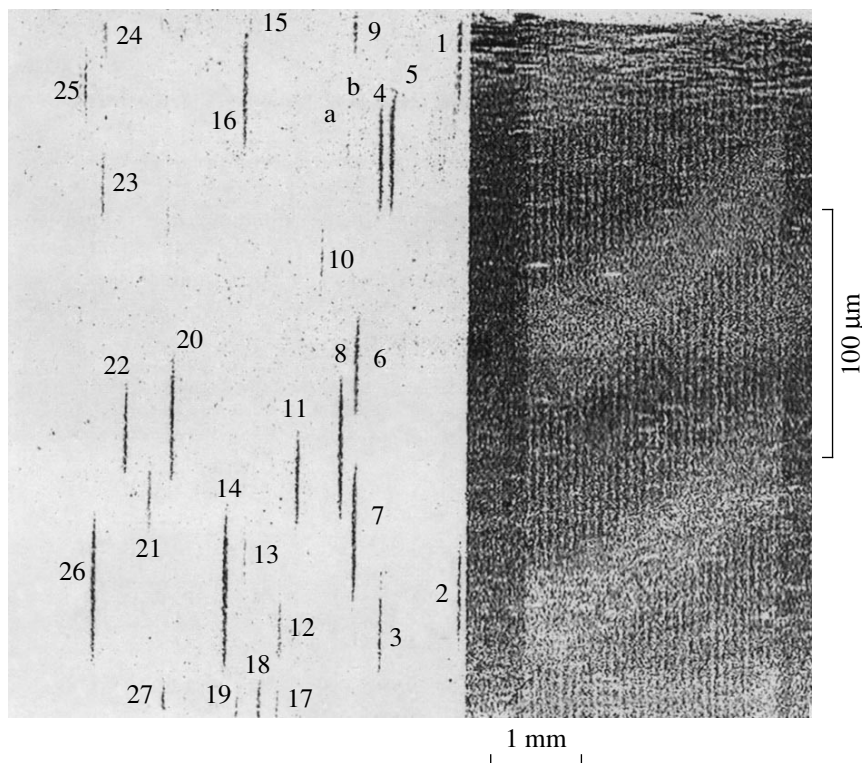

Fig. 4. Experimental X-ray image obtained for the objects with a 20-fold enlargement in one direction.

The results of our X-ray experiments are presented in Fig. 4. The image of the copper gauze is a system of sticks separated by light strips. One stick corresponds to one square hole. Considering that the enlargement was only in one direction, the ratio between the length and width of the stick corresponds to the actual enlargement and, as expected, equals ~20. The image of holes in the impact membranes also represents a system of
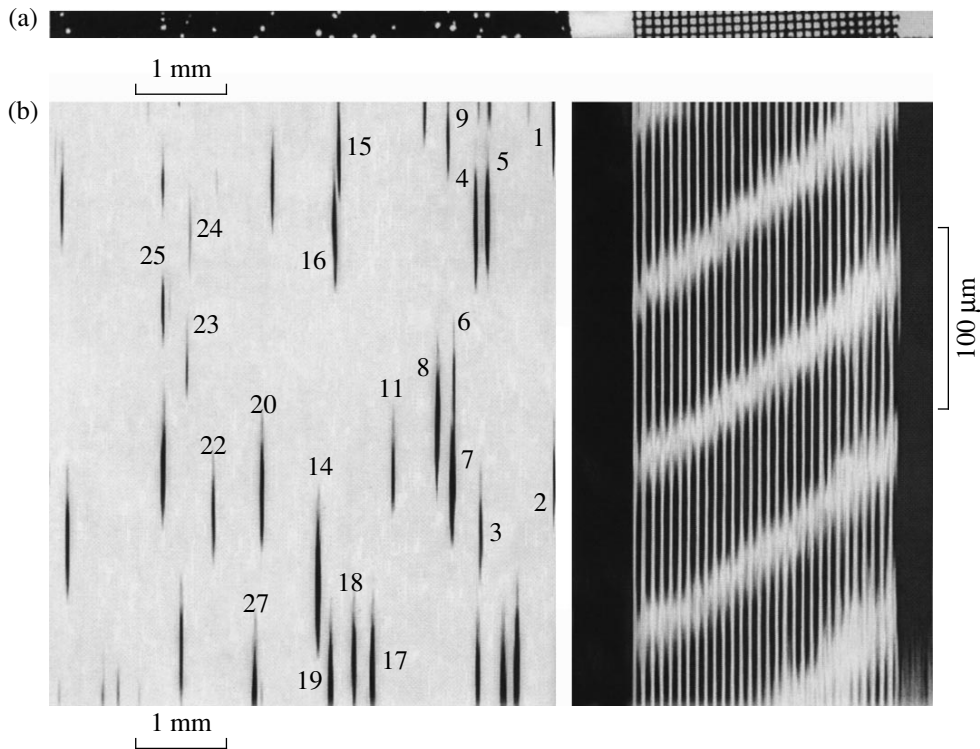
Fig. 5. Image of the same objects obtained on an optical microscope: (a) optical image and (b) the same image extended 20-fold in one direction.

sticks. This image can be juxtaposed with the image obtained on an optical microscope (Fig. 5). The pore diameters obtained in the X-ray experiment agree well with the electron microscopy data. The number of tracks distinguished by an optical microscope is smaller than their number found in the electron photomicrographs, indicating that not all tracks are through-going, although they can be distinguished in the X-ray image (Figs. 3, 4, 5, tracks 10, 12, 21).

Thus, we have substantiated the possibility of obtaining enlarged images of the objects with a resolution of ~20 µm. The images of some objects are obtained with a ~20-fold enlargement. The enlarged images are obtained for the holes in the impact track membranes with the smallest diameter being on the order of 20 µm.

REFERENCES

1. I. A. Artioukov, V. E. Asadchikov, V. I. Gulimova, *et al.*, in *Proceedings of the Sixth International Conference on X-ray Microscopy, Berkley, 1999*, p. 529.

2. O. P. Bratov, N. V. Denisov, I. P. Zhizhin, *et al.*, Apparatus and Methods in X-ray Analysis **4**, 3 (1969).

3. I. A. Artyukov, V. E. Asadchikov, A. I. Vilenskiĭ, *et al.*, Dokl. Akad. Nauk **372** (5), 608 (2000) [Dokl. Phys. **45**, 248 (2000)].

4. V. E. Asadchikov, V. I. Beloglazov, A. V. Vinogradov, *et al.*, Kristallografiya **44**, 592 (1999) [Crystallogr. Rep. **44**, 546 (1999)].

5. V. V. Aristov and A. I. Erko, *X-ray Optics* (Nauka, Moscow, 1991).

6. W. Yun, M. R. Howels, J. Feng, *et al.*, in *Proceedings of the Sixth International Conference on X-ray Microscopy, Berkley, 1999*, p. 529.

7. Y. Kagoshima, Y. Tsusaka, J. Matsui, *et al.*, in *Proceedings of the Sixth International Conference on X-ray Microscopy, Berkley, 1999*, p. 41.

Translated by V. Sakun

Spin Disclination in a Layered Antiferromagnet with a Screw Dislocation

B. A. Ivanov* and V. E. Kireev

Institute of Magnetism, National Academy of Sciences of Ukraine, Kiev, 03142 Ukraine
* e-mail: bivanov@i.com.ua
Received January 12, 2001

A screw dislocation perpendicular to layers in layered antiferromagnets with a ferromagnetic exchange interaction of spins in the atomic planes and an antiferromagnetic interaction between planes gives rise to nonsingular disclinations with a ferromagnetic core. © *2001 MAIK "Nauka/Interperiodica".*

PACS numbers: 61.72.Lk; 75.25.+z

Spin ordering in an antiferromagnet is described within the concept of a finite number of magnetic sublattices each ferromagnetically ordered, provided that the total magnetic moment of the antiferromagnet equals zero in the exchange approximation. Antiferromagnetic ordering is sensitive to crystal lattice defects, which destroy the sublattice structure of a perfect antiferromagnet. Lattice defects may give rise to an inhomogeneous spin distribution [1, 2].

Historically, the first example of such inhomogeneities was proposed by Dzyaloshinski [3] and Kovalev and Kosevich [4], who noted that the presence of an edge dislocation in an antiferromagnet brings about a "failure" in sublattices and gives rise to macroscopic magnetic defects, namely, domain walls and disclinations. Such effects arise in the case when the interface between the antiferromagnet and ferromagnet (FM) contains an atomic step (see [5, 6] and references therein). A common property of all these inhomogeneities is the appearance of antiferromagnetic disclinations that can be considered antiferromagnetic vortices with a half-integer value of the topological charge (vorticity) (see [1, 2]). An analysis of disclinations is of its own interest within the context of enhanced interest to two-dimensional magnetic solitons and especially to vortices (see [7–10]).

An edge dislocation with staggered ordering in an antiferromagnet was considered in all papers mentioned above. The occurrence of a jump in the antiferromagnetic vector **l** at a certain surface that rests on the dislocation line and extends to the crystal surface is a common property of all disclinations in an antiferromagnet with a dislocation in the lattice. However, the specific spin density distribution for various types of antiferromagnets may strongly differ from the distribution characteristic of the antiferromagnetic disclination described above. To demonstrate this fact, we consider a spin inhomogeneity that arises in layered antiferromagnets like $CoCl_2$, $FeCl_2$, and $NiCl_2$ with screw dislocations perpendicular to the layers. In the antiferromagnets indicated above, the spin-exchange interaction in the planes is ferromagnetic (exchange integral J), and the neighboring planes are coupled by an antiferromagnetic interaction with an exchange integral J', which is commonly weaker than the in-plane interaction. The typical ratio $J/J' \approx 10$–100, and it reaches 10^4–10^6 for intercalated systems [11]. We will show that the smallness of J'/J gives rise to nonsingular disclinations, which can be consistently described within a continual theory.

The Hamiltonian of a perfect magnet of the type indicated above can be written as

$$\mathcal{H} = \sum_{n=1}^{N_z} \sum_{\mathbf{r},\boldsymbol{\delta}} [-J\mathbf{S}_{n,\mathbf{r}}\mathbf{S}_{n,\mathbf{r}+\boldsymbol{\delta}} \\ + J'\mathbf{S}_{n,\mathbf{r}}\mathbf{S}_{n+1,\mathbf{r}+\boldsymbol{\delta}} + K(S_{n,\mathbf{r}}^z)^2], \qquad (1)$$

where n numbers the layers with a ferromagnetic interaction (exchange integral J), **r** determines the sites in one layer, $\boldsymbol{\delta}$ designates the vectors from the nearest neighbors to a given site inside the layer, $J' \ll J$ is the exchange coupling between the layers, and $N_z = L_z/a_z$ is the number of atomic planes. We also took into account easy-plane anisotropy with a constant K.

The summation in the Hamiltonian given by Eq. (1) is carried out over spins inside each layer and, then, over individual layers. If a screw dislocation is introduced into the crystal, the same classification can be made for any crystal region that does not contain the dislocation line, but the resulting system of atomic planes globally changes to a screw surface (see [12] for more detail and figure). It is convenient to parametrize this surface by coordinates r and χ, where χ varies continuously from 0 to $2\pi N_z$. With the use of these coordinates, one can proceed to the continual limit for the spin density $\mathbf{S} = S\mathbf{m}$, $\mathbf{m}^2 = 1$. Let us use the long-wave

approximation for spins lying locally in the same plane [changing, for example, $\mathbf{S}(\mathbf{r} + a\mathbf{e}_x) + \mathbf{S}(\mathbf{r} - a\mathbf{e}_x)$ to $2\mathbf{S}(r) + a^2\partial^2\mathbf{S}/\partial x^2$, and so forth] and retain the discrete description for neighboring planes. Then, the following equation can be obtained for the macroscopic magnet energy:

$$\mathcal{W} = JS^2\int_0^R rdr \int_0^{2\pi L_z/a_z} d\chi\left\{\frac{1}{2}(\nabla\mathbf{m})^2\right.$$

$$+ \frac{j}{a^2}\mathbf{m}(r, \chi)[\mathbf{m}(r, \chi + 2\pi) \qquad (2)$$

$$\left. + \mathbf{m}(r, \chi - 2\pi)] + \frac{K}{Ja^2}m_z^2(r, \chi)\right\}.$$

Here, the gradient is taken only in the given plane, $j = J'/J$, the quantities $\mathbf{m}(r, \chi)$ and $\mathbf{m}(r, \chi + \pi)$ correspond to spins arranged in neighboring planes, and the z axis is perpendicular to the layers. For simplicity, we assume that the sample is cylindrical in shape and has radius R.

Let us start with an analysis of the simple case of the *XY* model, in which the easy-plane anisotropy is extremely strong and all spins lie in the *xy* plane; that is, $m_z = 0$, $m_x = \cos\phi$, and $m_y = \sin\phi$. In the case of the *XY* model, the problem can be analyzed exactly. In particular, the occurrence of the second-order phase transition homogeneous state–spin disclination can be demonstrated. Next, we will show that a nonsingular disclination with a ferromagnetic core may arise in the case when the anisotropy is finite and spins can be offset from the *XY* plane.
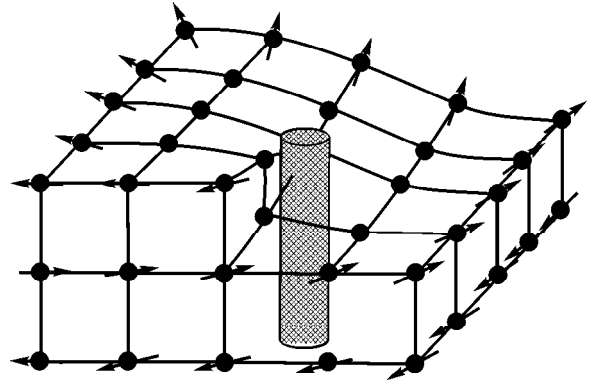
For the *XY* model, energy (2) takes the form

$$\mathcal{W} = JS^2\int_0^R rdr \int_0^{2\pi L_z/a_z} d\chi\left\{\frac{1}{2}(\nabla\phi)^2\right.$$

$$+ \frac{j}{a^2}[\cos[\phi(\chi) - \phi(\chi + 2\pi)] \qquad (3)$$

$$\left. + \cos[\phi(\chi) - \phi(\chi - 2\pi)] + 2]\right\}.$$

The extremals of the functional in Eq. (3) are defined by the equation

$$\nabla^2\phi + \frac{j}{2a^2}\sin[\phi(\chi) - \phi(\chi + 2\pi)]$$

$$+ \frac{j}{2a^2}\sin[\phi(\chi) - \phi(\chi - 2\pi)] = 0. \qquad (4)$$

It is easy to verify that a regular rotation of the type

$$\phi = \frac{\alpha\chi}{2} + \phi_0 \qquad (5)$$

Screw dislocation in a magnet (the core dislocation region is conventionally designated by a cylinder) and the spin distribution (arrows) in the spiral line enclosing the dislocation line.

corresponds to an exact solution of this equation. Note that α can correspondingly take arbitrary values, as distinct from the case of a vortex or a common disclination where α is an integer or half-integer value. This behavior of α is determined by the properties of the Riemann surface at which the variable ϕ is defined, namely, by the fact that there is no need to introduce the continuity conditions for $\cos\phi$ and $\sin\phi$ when χ is rotated through 2π.

Substituting the solution given by Eq. (5) into Eq. (3) and taking into account that, by virtue of Eq. (5), $(\nabla\phi)^2 = \alpha^2/4r^2$, we obtain the magnet energy as a function of parameter α, J'/J ratio, and system size

$$\mathcal{W} = \pi JS^2\frac{L_z}{a}\left[\frac{\alpha^2}{4}\ln\left(\frac{R}{r_0}\right) + \frac{jR^2}{a^2}(1 + \cos\pi\alpha)\right], \quad (6)$$

where r_0 is the cutoff radius on the order of the lattice constant in the layer.

From the energy minimum condition $d\mathcal{W}/d\alpha$, it follows that two types of magnet states are possible: a homogeneous state with $\alpha = 0$ and inhomogeneous states that correspond to the values of α obeying the transcendent equation

$$\frac{\sin\pi\alpha}{\pi\alpha} = \frac{\ln(R/r_0)}{2\pi^2 j(R/a)^2}. \qquad (7)$$

The homogeneous (ferromagnetic) state is unfavorable from the point of view of the antiferromagnetic coupling between the layers. It corresponds to the magnet energy $\mathcal{W}_{hom} = 2J'S^2V/v_0$, where $V = \pi R^2 L_z$ is the magnet volume, and $v_0 = a^2a_z$ is the unit cell volume. The inhomogeneous solutions exist only for large particles at $R > R_c$, where R_c is defined by the equation

$$2\pi^2 j\left(\frac{R_c}{a}\right)^2 = \ln\frac{R_c}{r_0}. \qquad (8)$$

At small $j \ll 1$, the value of R_c is macroscopic, $R_c \gg a$. Therefore, it is reasonable to consider the transition

from an inhomogeneous state, which can exist at $R > R_c$, to the homogeneous state, which is the only possible one at $R \leq R_c$. Here, α serves as the order parameter, and $\alpha \propto \sqrt{(R - R_c)/R_c}$ in the vicinity of the transition point. If $R \gg R_c$, the term on the right-hand side of Eq. (7) is small, the value of α is close to unity, $\alpha = 1 - (R_c/R)^2$, so that the spin distribution is almost identical to the distribution of vector $\mathbf{l}$ in an antiferromagnetic disclination, where $\phi = \chi/2 + \phi_0$, and the energy in Eq. (3) is defined by the equation

$$E = \frac{\pi J L_z}{4a} \ln\left(\frac{R}{r_0}\right) \ll J N_{at} \ll \mathscr{W}_{hom}. \qquad (9)$$

In this case, spins in neighboring layers are almost anti-parallel, and we may pass to the description in terms of the antiferromagnetic vector $\mathbf{l}$, which is defined for layered antiferromagnets as the difference of spins located one above the other. It is clear that such a vector can be introduced locally for a region that does not contain the dislocation line. The antiferromagnetic vector $\mathbf{l}$ defined in this way will have a discontinuity at a certain surface that rests on the dislocation line and extends to the crystal surface.

Taking into account the real three-dimensional nature of spins, which is essential at $K \ll J$ (at any ratio between K and J' under the condition that $J' \ll J$), leads to much more complicated equations. Therefore, we will restrict our consideration to the case of large R. In this case, the energy in Eq. (3) is conveniently expressed in terms of common angular variables for the unit spin vector $S_z = S\cos\theta$, $S_x = S\sin\theta\cos\phi$, $S_y = S\sin\theta\sin\phi$. The energy in this parametrization takes the form

$$\mathscr{W} = J S^2 \int_0^R r\,dr \int_0^{2\pi L_z/a_z} d\chi \left\{ \frac{1}{2}[(\nabla\theta)^2 + (\nabla\phi)^2\sin^2\theta] \right.$$

$$+ \frac{1}{2\Delta_0^2}\cos^2\theta + \frac{j}{a^2}[\cos\theta(\cos\theta_+ + \cos\theta_-) \qquad (10)$$

$$\left. + \sin\theta\sin\theta_+\cos(\phi - \phi_+) + \sin\theta\sin\theta_-\cos(\phi - \phi_-)] \right\}.$$

Here, we introduced parameter Δ_0 with the dimensionality of length, $\Delta_0^2 = Ja^2/K$, and notations $\theta_\pm = \theta(r, \chi \pm 2\pi)$ and $\phi_\pm = \phi(r, \chi \pm 2\pi)$. It is easy to verify that the system of Euler–Lagrange equations for the functional given by Eq. (10) has a solution of the form

$$\theta = \theta_0(r), \quad \phi = \frac{\alpha\chi}{2} + \phi_0. \qquad (11)$$

Here, the function θ obeys an ordinary differential equation of the same type as for vortices in magnets, see [1, 2],

$$\frac{d^2\theta}{dr^2} + \frac{1}{r}\frac{d\theta}{dr} + \sin\theta\cos\theta\left(\frac{1}{\Delta_\alpha^2} - \frac{\alpha^2}{4r^2}\right) = 0, \qquad (12)$$

with the only distinctions that α can now be arbitrary rather than integer, and that the characteristic length Δ_α depends on the parameter α

$$\Delta_\alpha^2 = \frac{Ja^2}{K + 2J'\sin^2(\pi\alpha/2)} < \Delta_0^2. \qquad (13)$$

The natural boundary conditions for this equation are determined by the absence of singularities at the center ($\theta = 0$ or π at $r = 0$) and by the fact that, at large distances from the disclination center, spins lie in the easy xy plane, forming the same structure as that considered above for a disclination in the XY model,

$$\cos\theta = p \quad \text{at} \quad r \longrightarrow 0, \quad p = \pm 1,$$
$$\theta \longrightarrow \pi/2 \quad \text{at} \quad r \longrightarrow \infty. \qquad (14)$$

Thus, the additional discrete parameter p arising in the problem plays the same role as the polarization for the out-of-plane vortex in easy-plane ferromagnets (see [2]). The disclination state is degenerate not only in the sign of α but also in the value of p.

The solution of Eq. (12) at $r \longrightarrow 0$ can easily be constructed numerically with regard to Eq. (14) and asymptotic behavior $\sin\theta \simeq (r/\Delta_\alpha)^{\alpha/2}$. For large R, $\alpha \simeq 1$ and the disclination energy is

$$E = \frac{\pi J S^2}{4}\ln\left(\frac{R}{C\Delta}\right), \quad \Delta = a\sqrt{\frac{J}{K + 2J'}}. \qquad (15)$$

Thus, this disclination has a typical logarithmic dependence on the sample size. However, this is the only similarity of a disclination in a layered antiferromagnet to the disclination in an antiferromagnet with a staggered structure considered in the previous article. Let us list their distinctions.

It is pertinent to refer to a disclination in a layered antiferromagnet as ferromagnetic, because the ordering of spins in the core is ferromagnetic along the entire length of the disclination line and the arrangement of spins in each of the local regions of atomic planes is close to ferromagnetic. The total magnetization in the basal plane of a magnetic particle composed of an odd number of atomic planes is macroscopic and close to $(1/2)2\mu_B S N_{at}$, where N_{at} is the number of atoms in the plane.

If reasonable conditions $K \ll J$ and $J' \ll J$ are fulfilled, the core of a ferromagnetic disclination contains no singularity and the cutoff parameter in the energy takes a macroscopic value $\Delta \gg a$.

The state of a ferromagnetic disclination is degenerate with respect to two discrete numbers, p and the sign

of α, rather than one, as in a singular antiferromagnetic disclination. This feature can manifest itself in the effects of macroscopic quantum tunneling for spin disclinations, see [13].

REFERENCES

1. A. M. Kosevich, B. A. Ivanov, and A. S. Kovalev, *Nonlinear Magnetization Waves: Dynamic and Topological Solitons* (Naukova Dumka, Kiev, 1983).

2. V. G. Bar'yakhtar and B. A. Ivanov, Sov. Sci. Rev., Sect. A **16**, 1 (1992).

3. I. E. Dzyaloshinskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **25**, 110 (1977) [JETP Lett. **25**, 98 (1977)].

4. A. S. Kovalev and A. M. Kosevich, Fiz. Nizk. Temp. **3**, 259 (1977) [Sov. J. Low Temp. Phys. **3**, 125 (1977)].

5. R. L. Stamps and R. E. Camley, Phys. Rev. B **54**, 15200 (1996); R. L. Stamps, R. E. Camley, and R. J. Hicken, Phys. Rev. B **54**, 4159 (1996).

6. O. K. Dudko and A. S. Kovalev, Fiz. Nizk. Temp. **25**, 25 (1999) [Low Temp. Phys. **25**, 18 (1999)].

7. A. S. Kovalev, Fiz. Nizk. Temp. **20**, 1034 (1994) [Low Temp. Phys. **20**, 815 (1994)].

8. B. A. Ivanov, V. E. Kireev, and V. P. Voronov, Fiz. Nizk. Temp. **23**, 845 (1997) [Low Temp. Phys. **23**, 635 (1997)].

9. O. K. Dudko and A. S. Kovalev, Fiz. Nizk. Temp. **24**, 559 (1998) [Low Temp. Phys. **24**, 422 (1998)].

10. O. K. Dudko and A. S. Kovalev, Fiz. Nizk. Temp. **26**, 821 (2000) [Low Temp. Phys. **26**, 603 (2000)].

11. D. C. Wiesler, H. Zabel, and S. M. Shapiro, Z. Phys. B **93**, 277 (1994).

12. A. M. Kosevich, *Dislocations in the Theory of Elasticity* (Naukova Dumka, Kiev, 1978).

13. B. A. Ivanov and A. K. Kolezhuk, in *Frontiers in Magnetism of Reduced Dimension Systems*, Ed. by V. G. Bar'yakhtar, P. E. Wigen, and N. A. Lesnik (Kluwert, Dordrecht, 1998), NATO ASI Ser., Ser. 3, **49** (1998).

Translated by A. Bagatur'yants

Observation of Laser-Induced Local Modification of Magnetic Order in Transition Metal Layers

A. M. Alekseev[1], Yu. K. Verevkin[2], N. V. Vostokov[3], V. N. Petryakov[2], N. I. Polushkin[3,4], A. F. Popkov[1], and N. N. Salashchenko[3]

[1] *Lukin Scientific Research Institute of Physical Problems, Moscow, 103460 Russia*
[2] *Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603600 Russia*
[3] *Institute of Macrostructure Physics, Russian Academy of Sciences, Nizhni Novgorod, 603600 Russia*
[4] *e-mail: nip@ipm.sci-nnov.ru*
Received January 15, 2001

Laser-induced local modifications of magnetic order in thin Fe–Cr layers were investigated. Local modification in the layers were induced by interfering laser beams. The results of the study give evidence for the formation of submicron-sized anisotropically shaped ferromagnetic regions with a well-defined direction of the easy magnetic axis in the interference maxima at the modification threshold. It was also found that the magnetic anisotropy of a medium is drastically reduced with changing the shapes of these local regions and distances between them. This may be due to the strengthening of the interaction between the regions through the paramagnetic matrix. © *2001 MAIK "Nauka/Interperiodica"*.

PACS numbers: 75.70.Kw; 75.60.Jk; 75.50.Bb; 75.50.Cc

It is known [1–3] that the Fe(Co)-based nanocomposite alloys of 3*d* metals may not possess long-range magnetic order at the concentrations of magnetic component as high as 70–75 at. %. In [1], a model was suggested according to which such alloys are composed of superparamagnetic clusters in a nonmagnetic medium. As the concentration of magnetic atoms increases, the cluster formation becomes more intense and, eventually, a concentration is achieved (percolation threshold) at which the long-range magnetic order is extended over the whole material.

Another way of creating a magnetically ordered state in alloys (without changing the mean concentration of the magnetic component) consists in heating a metallic medium by an intense (~1 J/cm^2) short (~10 ns) laser pulse to a temperature higher than the melting point, to initiate diffusional mixing of the alloy components, and its fast cooling (with a rate of ~10^{10} K/s) after the pulse. This gives rise to a metastable supersaturated solid solution whose magnetic properties are sensitive to the positional and chemical short-range orders in the system [4]. In [5, 6], we reported the observation of thermo- and laser-induced modifications of the magnetic order in thin-film (< 100 nm) mixtures of the Fe−C, Co–C, and Fe–Cr types. The major experimental fact was that after the laser pulse with an energy density of 200–400 mJ/cm^2, the originally (super)paramagnetic medium became ferromagnetic at room temperature in a certain (rather narrow) concentration range of the magnetic component, with the saturation magnetization of the resulting ferromagnet being close to the value in the bulk Fe and Co samples.

In this work, laser-induced magnetic ordering was used to produce a regular grating of small ferromagnetic elements. The sizes of these elements are limited in all three directions and comparable with the key micromagnetic parameters: exchange length (10 nm) and domain wall thickness (10–100 nm). As regards the physics of magnetization reversal, these elements are intermediate between the multidomain and single-domain systems. The physical behavior of this system is noteworthy in two aspects. First, it is important to determine the conditions for the formation of single-domain (ferro)magnetic regions in the course of local modification of a paramagnetic medium. Second, it is of interest to study the specific features of magnetization reversal for a system of magnetically hard particles in a magnetically soft medium. Interest in the structures of small magnetic objects has arisen because such systems are viewed as an alternative material for the design of new magnetic ultrahigh-density recording and data storage devices [7, 8].

In this work, local modification was accomplished by the coherent UV laser beams that are capable of creating an interference grating with submicron spacing (down to 200 nm) at the sample surface. The sizes of modified magnetic elements can preliminary be estimated from the thermal diffusion length $\sqrt{a\tau}$, where a is the thermal conductivity coefficient and τ is the heat pulse duration. For a pulse duration of no more than several nanoseconds and a pulse energy close to the magnetic transformation threshold, the sizes of the elements can be expected not to exceed 100 nm.
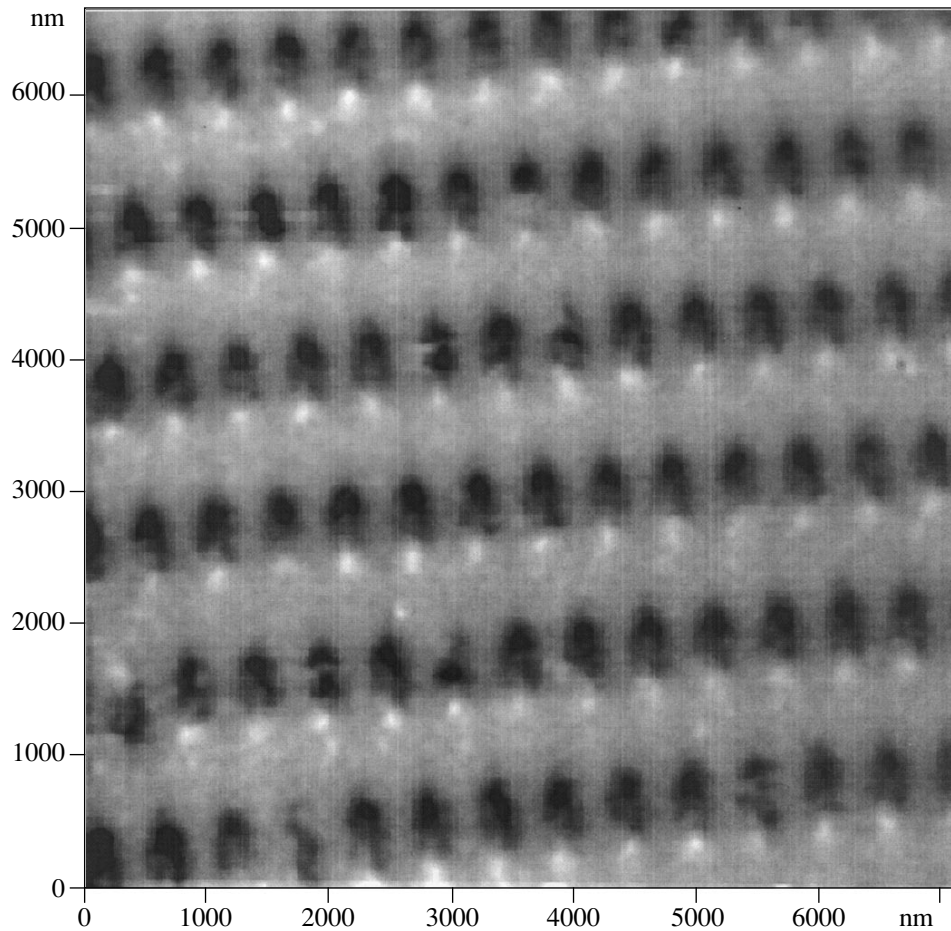
Fig. 1. The MFM image of the surface of a $Fe_{0.7}Cr_{0.3}$ layer irradiated by two pairs of beams with energy density $E = 250$ mJ/cm^2. The asymmetrically shaped dipoles are clearly seen. The image is obtained for a residual state after the magnetization along the dipole direction.

The starting alloys for the subsequent laser annealing were thin layers (15–20 nm) prepared by the alternate deposition of small portions (0.3–0.5 nm) of Fe(Co) and Cr(C) on silicon substrates, with the Nd^{3+} laser radiation being focused onto the targets placed in a vacuum chamber. The absence of ferromagnetic order in the starting samples and its appearance after irradiation was monitored by the spectra of ferromagnetic resonance. These data were used to determine the optimum concentration of the magnetic component in a mixture for which the starting sample did not yet show a magnetic response at room temperature while, after irradiation, a well-defined signal of ferromagnetic resonance appeared.

The magnetic structures were produced using a narrow-band (0.04 cm^{-1}) XeCl eximer laser ($\lambda = 308$ nm) with a pulse energy up to 50 mJ and a pulse duration of 8 ns. Laser emission was monochromatized by the intracavity mode selection using a Fabry–Pérot interferometer to provide a contrast interference pattern throughout the whole cross section of the laser beam. To produce a two-dimensional grating of modified local regions, the laser radiation was split into two pairs of beams that were incident on the sample at different angles in two mutually perpendicular planes. The intensity distribution in the interference maxima was extended along the axis in the plane of the smaller angle of incidence, with the aspect ratio corresponding to the ratio of angles of incidence. The total area of the array was determined from the diffraction of a HeNe laser radiation by the grating. Depending on the pulse energy, this area was 5–10 mm^2.

The properties of the gratings were studied by atomic-force (AFM) and magnetic-force (MFM) microscopy, and the magnetic hysteresis loops were measured in the presence of the longitudinal magneto-optical Kerr effect. The AFM/MFM data were recorded on a Solver P47 (NT-MDT, Moscow) scanning probe microscope. Silicon cantilevers sputtered with a ≈30-nm thick Co layer were used as magnetic probes. Resonance frequencies of the probes were equal to 55–120 kHz. Before measurements, the magnetic probe was magnetized along the tip axis while the samples themselves were preliminary magnetized in fields
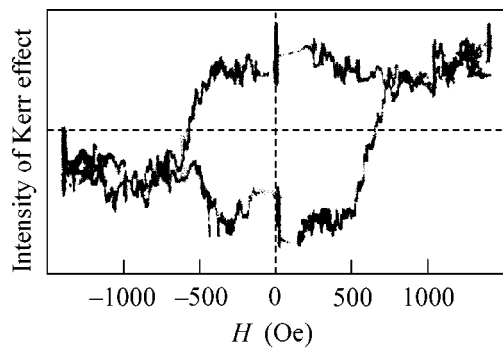
Fig. 2. Magnetization reversal curve for the magnetic elements in a $Fe_{0.7}Cr_{0.3}$ layer subjected to the interference laser annealing ($E = 250$ mJ/cm^2). The external magnetic field H is applied in the dipole direction.

of up to 1500 Oe aligned in the sample plane with the minor or major axis of the created elements. The distribution of the magnetic force gradient was determined from the phase shift induced in the cantilever oscillations by either attractive or repulsive forces acting on the magnetic tip.

Figure 1 is the MFM image of the surface of a $Fe_{0.7}Cr_{0.3}$ layer after interference irradiation at angles of incidence of 10° and 40° by a laser pulse with an energy density of $E = 250$ mJ/cm^2. This energy is close to the very threshold of magnetic modification (220–240 mJ/cm^2). It has not yet produced any changes in the surface relief. Nevertheless, a periodic structure of identically oriented asymmetrically shaped dipoles is clearly seen in the MFM image: dark (large) and light (smaller) spots indicate the poles of the magnetic elements. The observed orientation of the dipoles corresponds to the sample premagnetization direction. One can see from the pole positions that the aspect ratio of the ferromagnetic elements equals approximately 1 : 4, in quantitative agreement with the ratio of angles of incidence.

Figure 2 shows the magnetization reversal curve for the same Fe–Cr sample. In this experiment, an external magnetic field was aligned with the dipole orientations in Fig. 1. One can see that this curve has a hysteretic character typical of ferromagnets, and the shape of the hysteresis loop is close to rectangular with a magnetization reversal field of ≈500 Oe. An attempt at reversing magnetization of this sample in the direction perpendicular to the dipole orientation in fields up to 2000 Oe did not affect the magneto-optical response to within the noise level. Therefore, the irradiation of the layer near the magnetic transformation threshold gives rise to a magnetic structure with a well-defined easy magnetization direction corresponding to the orientation of dipoles in the MFM image.

A characteristic feature of the magnetic elements is that their thickness (15–20 nm) is much smaller than their lateral sizes (100 × 400 nm). We carried out

micromagnetic modeling of magnetization distribution in the elements with thicknesses 10–20 nm. It follows from the numerical experiments that the magnetic elements show a single-domain behavior at thicknesses less than 12 nm in the case of weak anisotropy or at 15 nm in the case of a uniaxial anisotropy of ~10^6 erg/cm^3. Otherwise, skew-symmetric magnetization distributions with edge spin fixation or a more complex spin configuration with one or two vortices inside the elements are formed. Hence, the magnetic-pole asymmetry observed in our experiments (Fig. 1) may point to the inhomogeneous magnetization distribution in the magnetic elements.

It is of interest to examine the properties of a layer modified with higher-energy laser pulses. Figure 3 is the AFM image of the topography of a surface Fe–Cr layer irradiated by $E = 300$ mJ/cm^2. This image provides evidence for the formation of craters in the interference maxima as a result of melting and expelling melt toward the periphery of the interference maxima by vapor pressure. The craters are shaped like ellipses with an axes ratio of 1 : 2, and their depth from the bottom to the top of their rims is as large as 8–10 nm. The modification of surface topography is also accompanied by profound changes in the properties of the created magnetic structures. The magnetic hysteresis loops of the same ($E = 300$ mJ/cm^2) Fe–Cr sample suggest that magnetic anisotropy in the new matrix is strongly reduced (Fig. 4). It is seen that the remanent magnetization of this sample remains close to the saturation magnetization in an external magnetic field oriented along both the major (Y) and the minor (X) crater axes. Nevertheless, it is worth noting that the magnetization reversal loop becomes more rectangular for remagnetization along the X axis (minor axis of the craters in Fig. 3), possibly because of the appearance of the easy magnetic axis in this direction, whereas the magnetization in a "hard" Y direction shows a relatively smooth saturation. The shape of the hysteresis loop in this direction is typical of a strongly dispersive medium [9]. Note that if the system of magnetic elements is well-defined in the topographical images, then the anomalous orientation of the easy magnetic axis occurs even for loose packing of elliptic craters. As the laser energy further increases (350–400 mJ/cm^2), the minor axis of the ellipses increases up to a contact of two neighboring crater rims. However, these relief modifications no longer alter the magnetic behavior of the irradiated samples.

The studies of the morphology and magnetic properties of the Fe–Cr-type systems suggest [2, 3] that they consist of Fe-enriched superparamagnetic grains no larger than several nanometers. According to the corresponding phase diagram [10], the Fe and Cr components form unlimited solutions at temperatures higher than the melting point (≈1600 K), while the supersaturated Fe_xCr_{1-x} ($x = 50$–70 at. %) solution possesses a higher Curie temperature (≈900 K). The necessity of
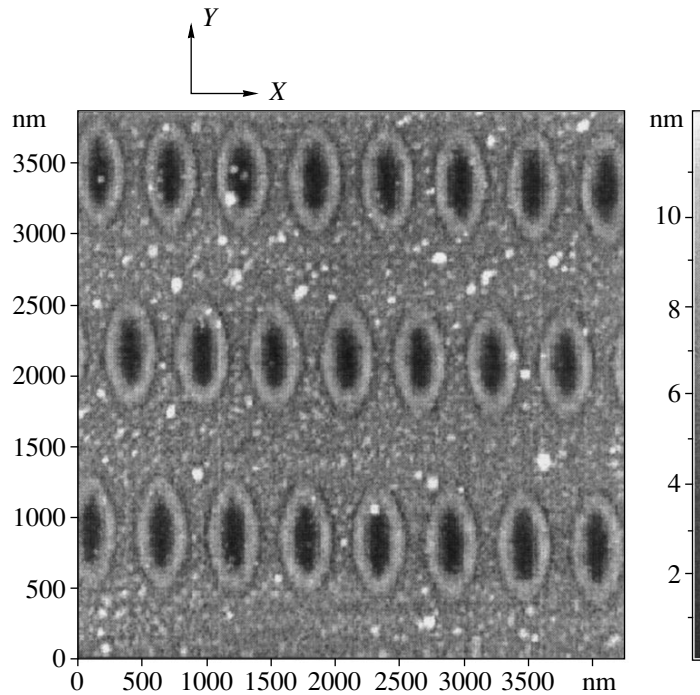
Fig. 3. Topography of the surface of a $Fe_{0.7}Cr_{0.3}$ layer after the irradiation with $E = 300$ mJ/cm^2. The AFM image.

achieving the liquid state upon laser heating in an attempt to accomplish magnetic transformation is corroborated by the fact that the melting and expelling of a liquid from the interference maxima (Fig. 3) are observed for a laser energy (300 mJ/cm^2) only slightly above the magnetic transformation threshold (220–240 mJ/cm^2). However, the most pronounced and well-interpreted magnetic structure is formed directly at the modification threshold (Figs. 1, 2), where the craters are as yet not formed at the surface.

It is not improbable that the observed increase in the "rectangularity" of the magnetization loops of the crater structures (Fig. 4) in the transverse direction can be explained by the interactions between the ferromagnetic regions (craters) via the superparamagnetic medium. This interaction becomes possible either because of the medium magnetization by the stray fields of the ferromagnetic regions or due to medium modification upon an increase in the cluster sizes and cluster approach. Evidently, the separation between craters and their aspect ratio decrease with increasing pulse energy. Since the distance between the elliptic craters along their minor axes is appreciably shorter than along the major axes (Fig. 3), the demagnetization effect in this direction is less pronounced. Thus, the direction of the easy magnetic axis in the resulting crater structures is likely determined by the competition of two factors: the interaction between the ferromagnetic regions via a magnetically soft (superparamagnetic) medium and a change in the shapes of these regions.

Note in conclusion that our experiments can be used to formulate the technological requirements on the variable parameters (concentration of the magnetic component, laser pulse energy and duration, angles of incidence, etc.) for producing gratings of magnetic single-domain elements with spacing in the far-submicron
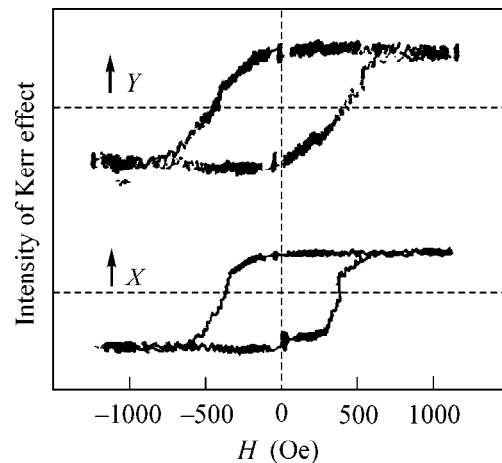

Fig. 4. Magnetization reversal curves for a $Fe_{0.7}Cr_{0.3}$ layer with the surface lattice of elliptically shaped craters ($E = 300$ mJ/cm^2). The loops are obtained for different orientations of H: along the major (Y) and minor (X) axes of ellipses.

range. The small sizes of the elements facilitate the realization of practically important physical properties (single-domain structure, rectangular hysteresis loops, short magnetization-reversal time, etc. [11]) for small magnetic objects.

REFERENCES

1. C. Schwink, K. Emmerich, and U. Schulze, Z. Phys. B **31**, 385 (1978).

2. K. Takanashi, T. Sugawara, K. Hono, and H. Fujimori, J. Appl. Phys. **76**, 6790 (1994).

3. M. Ohkoshi, J. Appl. Phys. **63**, 2926 (1988).

4. K. Handrich and S. Kobe, *Amorphe Ferro- und Ferrimagnetika* (Akademie-Verlag, Berlin, 1980; Mir, Moscow, 1982).

5. N. I. Polushkin and N. N. Salashchenko, J. Magn. Magn. Mater. **124**, 347 (1993).

6. Yu. Blyakhman, N. I. Polushkin, A. D. Akhsakhalyan, *et al.*, Phys. Rev. B **52**, 10303 (1995).

7. S. Y. Chou, Proc. IEEE **85**, 652 (1997).

8. R. M. H. New, R. F. W. Pease, and R. L. White, J. Magn. Magn. Mater. **155**, 140 (1996).

9. J.-G. Zhu and H. N. Bertram, J. Appl. Phys. **66**, 1291 (1989).

10. O. Kubaschewski, *Iron-binary Phase Diagrams* (Springer-Verlag, Berlin, 1982; Metallurgiya, Moscow, 1985).

11. D. D. Awschalom and D. P. DiVincenzo, Phys. Today **48**, 43 (1995).

Translated by V. Sakun

Elastic Properties of Substances in the Megabar Pressure Range: Inversion of Shear Rigidity

V. V. Brazhkin and A. G. Lyapin

Institute for High-Pressure Physics, Russian Academy of Sciences, Troitsk, Moscow region, 142190 Russia
Received January 25, 2001

The behavior of elastic moduli of substances is analyzed in the megabar pressure range. A new effect—inversion of the shear moduli and mechanical properties upon compression—is predicted for various classes of substances. The melting-curve data for different materials confirm the predicted phenomenon. The materials traditionally considered the softest, such as rare gas solids and molecular substances, may become the hardest in the megabar range. This should be taken into account in developing the experimental high-pressure technique. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 62.20.Dc; 62.50.+p

1. Investigations into the behavior of elastic properties and mechanical characteristics of strongly compressed materials are of crucial importance both to fundamental physics and geology and in high-pressure technique. In the quantum pressure range $P > Z^{10/3}m^4e^{10}/\hbar^8 \sim 10^2$–$10^4$ Mbar corresponding to the compression ratio of several tens of factors, all properties of a substance, including its elastic moduli, can be thought to be monotonic and almost linear functions of atomic number [1]. It is known that in the experimentally attainable megabar pressure range corresponding to the compression ratio of several times the properties of substances are, generally, nonmonotonic functions of atomic number [1, 2]. At the same time, it is *a priori* unclear whether the materials will still be divided into hard and soft in the megabar range, and which classes of substances will have the highest elastic moduli, hardness, etc., at such pressures.

It should be emphasized that the determination, by both experiment and computer simulation, of the mechanical properties such as hardness and strength of the substances under pressure is a highly challenging task. It is known, however, that these properties are closely related to the shear rigidity of materials and, correspondingly, to their elastic moduli [3], which, in turn, are the uniquely defined physical characteristics of substances. Therefore, the purpose of this work was to analyze the behavior of elastic, primarily, the shear moduli of substances in the megabar pressure range on the basis of the available experimental data and model empirical approaches. We will restrict ourselves mainly to rare gas solids and metals, because these classes are precisely the ones for which the predicted effect is expected to be most pronounced.

2. At normal pressure, the bulk and shear moduli of various substances lie in the range from kilobars to megabars [4, 5]. On compression, the elastic moduli increase (except for the possible anomalous behavior near phase transitions). For most substances, the experimentally measured derivatives of the bulk and shear moduli with respect to a pressure range from 1 to 8 [4, 5]. Consequently, the elastic moduli at megabar pressures are more likely determined by the corresponding baric derivatives rather than by the initial values of the moduli. At pressures of 3 Mbar and higher, the elastic moduli of every substance are higher than those of diamond at normal pressure, the record holder among all materials. As a result, there are no soft, in the ordinary sense, substances in the megabar range.

The behavior of bulk moduli of the majority of substances has been much studied up to megabar pressures [1, 2]. By contrast, the shear moduli of the majority of substances are measured only up to pressures of several tens of kilobars [4, 5]. As to the megabar range, only indirect estimates are known for the behavior of elastic shear constants of some substances, including some metals (Fe, Au, and Mo) [6, 7].

At the same time, the behavior of elastic constants of various classes of substances at high pressures can be analyzed by using a simple model of a system of particles interacting via the power n/m-potential $U = A/r^n - B/r^m$. With an increase in pressure, the effective repulsion exponent n lying in the interval from 6 to 12 for the majority of substances becomes dominant. This is so because the relative contribution to the elastic moduli and pressure from the noncentral forces, energy of valence electrons, and interior Coulomb forces, as well as from the attractive forces (the $-B/r^m$ term in the n/m-potential) rapidly drops upon compression, as compared to the contribution from the central repulsive forces between the nearest neighbors [8].

In this paper, we restrict ourselves for brevity only to the analysis of an FCC lattice that corresponds to one of the close-packed atomic arrangements. For the ele-
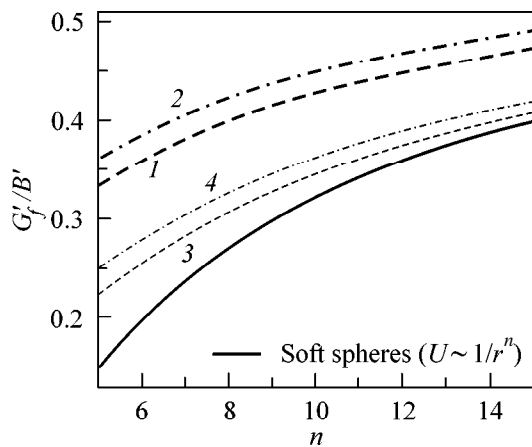
Fig. 1. The G'_f/B' ratio as a function of the repulsion exponent n for the FCC lattice with n/m-potential. (*1, 2*) $P = 0$ and $m/n =$ (*1*) 0.5 and (*2*) 0.8. (*3, 4*) Pressure corresponding to $V/V_0 = 0.5$ and $m/n =$ (*3*) 0.5 and (*4*) 0.8. The data for the soft-sphere model with the $1/r^n$ potential (see table) correspond to the n/m-potential ($m < n$) at $P \longrightarrow \infty$.

mentary cubic (simple cubic, BCC, and FCC) lattices with central interactions between atoms, one can readily obtain general expressions for the pressure and elastic constants [8, 9]:

$$P = \frac{1}{v}\sum_{R_i}[U'_{r^2}r_x^2]_{R_i}, \qquad (1)$$

$$c_{11} = \frac{2}{v}\sum_{R_i}[U''_{r^2}r_x^4]_{R_i} - P, \qquad (2)$$

$$c_{12} - P = c_{44} + P = \frac{2}{v}\sum_{R_i}[U''_{r^2}r_x^2r_y^2]_{R_i}, \qquad (3)$$

where v is the unit cell volume, the sums are over the lattice vectors R_i, and the derivatives of the interaction potential are taken with respect to the r^2 variable: $U'_{r^2} = dU(r^2)/d(r^2)$, etc. The values obtained in this work for the derivatives of the bulk and shear moduli of a lattice with purely repulsive potential $U = A/r^n$ in the short-range (soft spheres) and long-range interaction models are given in the table. Note that the bulk modulus of a lattice with central interaction is related to Voigt's shear modulus by the exact expression that follows from the Cauchy relation with pressure [10], $c_{12} = c_{44} + 2P$. Indeed, $G_f = (9c_{44} + 3(c_{11} - c_{12}))/15 = ((c_{11} + 2c_{12}) - 6P)/5 = (3B - 6P)/5$. The derivatives with respect to pressure are also obtained from these expressions:

$$G'_f = (3B' - 6)/5. \qquad (4)$$

It is significant that the derivatives of the elastic moduli and the relative shear elasticity of a lattice

$G/B \approx G'/B'$ both increase with n at ultrahigh pressures. The same trends hold in the general case of the n/m-potential (Fig. 1). With a rise in pressure, the derivatives of the elastic moduli of a lattice with n/m-potential tend to their soft-sphere ($U = A/r^n$) values (Fig. 2). For compressions $V/V_0 \sim 0.5$ (which correspond to real substances under pressures from tens of kilobars to the megabar [1]), they differ from the limiting values at $V/V_0 \longrightarrow 0$ by no more than 5–20%. That is, in the high-pressure limit ($P \longrightarrow \infty$), one has $G'_f(\infty)/B'(\infty) = 3/5[(n - 3)/(n + 3)]$. At the same time, it follows from Eq. (4) that, in a lattice with central forces (or, approximately, in any lattice at ultrahigh pressure), $\Delta(G'_f)/\Delta(B') = 3/5$ for any change in pressure. It then follows that the ratio $(\Delta(G'_f)/\Delta(B'))/(G'_f(\infty)/B'(\infty)) = (n + 3)/(n - 3)$ always exceeds unity and decreases with increasing n. Therefore, with a rise in pressure, the relative change in the derivative of the shear modulus is always larger than in the derivative of the bulk modulus, the effect being more pronounced with a decrease in the exponent n of the repulsive potential.

One can draw the following important conclusion from the above analysis. If the elastic moduli of one of any two substances are lower at normal conditions but its interatomic repulsive potential is more rigid (effective exponent n), then the ratios of moduli of these substances will be inverted under pressure and the final difference between their shear moduli will always be higher than between their bulk moduli. Evidently, the physical reason for such effects is associated with the negative contribution of pressure to the shear moduli [2, 9], in particular, to the $(c_{11} - c_{12})/2$ and c_{44} constants for the cubic lattice [see Eqs. (2) and (3)] and is formally clarified by Eq. (4).

3. Clearly, real substances are not described by the simple n/m-potential. However, one can draw from the above-mentioned analysis qualitative conclusions about the relative behavior of elastic properties of various classes of substances under strong compression. In particular, the effective interatomic potential in metals is softer than in rare gas solids. The effective repulsion exponent n in metals is smaller because of the screening effect of free valence electrons. One can naturally expect the inversion of shear moduli and, correspondingly, of other mechanical characteristics of these two classes of substances, with regard to the fact that rare gas solids are exceedingly soft substances at low pressures.

Note that simple estimates by Eqs. (1)–(3) yield quite realistic values for the derivatives of elastic moduli under pressure. For instance, the B' value in rare gas solids ($n \approx 12$ and $m \approx 6$) should change from 8 to 5 and G'_f should change from 3.6 to 1.8, in good agreement with experimental data [11, 12]. For transition metals, one can set $n \approx 8$ and $m \approx 1$ [13]. Accordingly, the B' value should change from 5 to 3.7 and G'_f from 1.8 to

1, also in good agreement with the experiment [4–7]. These estimates are also confirmed by other empirical calculations. The expression suggested in [5] gives $G' \sim 1.3 \pm 0.3$ for metals and $G' \sim 1.9 \pm 0.4$ for rare gas solids at megabar pressures. Lattice calculations with a more exact interaction model give $G' \approx 2$–2.5 for rare gas solids at $P \sim 0.5$ Mbar [2].

At a pressure of several megabars, the shear moduli of metals, probably become 1.5–2 times smaller than in rare gas solids (except, maybe, He), whereas the bulk moduli of metals become either close to or 10–50% higher than the rare-gas values. The fact that the B' and G' values in He, Li, LiH, and, probably, metallic hydrogen are smaller than in the other substances of their classes is caused by the absence of inner core electrons and by the soft repulsive potentials in these materials. At relatively low pressures, the derivatives in alkaline and some alkaline-earth metals are also small, $G' \sim 0.5$–1 [4, 5], and, accordingly, the shear moduli are several times lower than in other substances. However, after the electronic and phase transitions occurring in the pressure range from 50 kbar to 1 Mbar, the G' values in these substances should become close to the values in other metals.

4. The presently known empirical calculations and available experimental dependences do not contradict our hypothesis of the inversion of shear moduli in substances of various classes (Fig. 3a). Moreover, some indirect facts count in favor of this assumption. Within the framework of the Lindemann melting criterion, the baric behavior of the melting temperature T_m is primarily governed by the behavior of the shear modulus, as is also confirmed by experimental data [15].

As distinct from the shear moduli, the melting temperatures are experimentally measured or calculated *ab initio* for a series of substances of different classes up to pressures of several megabars. It turns out that the melting curves of metals in the megabar pressure range become flatter than those of nonmetals [16, 17]. The melting curves of inert substances such as Ar and of alkali-halide crystals such as NaCl intersect the melting curves of metals, e.g., Fe (Fig. 3b). The intersection of melting curves is a remarkable corroboration of the hypothesis of inversion of shear moduli of the corresponding classes of substances.

5. The results obtained in this work may find an important practical application to high-pressure technique in the problem of choosing between various substances as pressure-transmitting media. At room temperature and pressures higher than 120 kbar, all substances are solids. In particular, helium crystallizes and alcohol mixtures undergo glass transition. Nevertheless, the experimenters working in the megabar pressure range traditionally use solidified organic liquids or rare gas solids as working media at high pressures. In doing so, they groundlessly extend the conclusion about the softness of these substances to the megabar
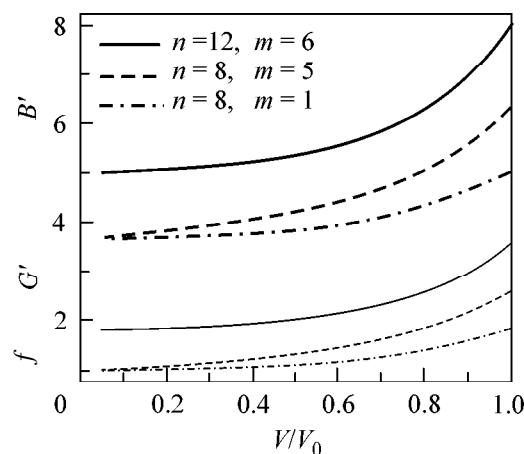

Fig. 2. The baric derivatives of bulk modulus (thick lines) and Voigt's shear modulus (thin lines) as functions of the compression ratio for the FCC lattice with the n/m-potentials typical of rare gases ($n = 12$ and $m = 6$) and d metals [13].

pressure range. As was pointed out above, no soft substances occur in the megabar range: the shear moduli of all materials are equal to several megabars and, accordingly, their flow stresses $\sigma_y \sim (0.05$–$0.1)G$ (see [18] and references cited therein) are equal to hundreds of kilobars. The ratio between the characteristics of different substances is an important parameter.

Pd^3/U_a	$2n\sqrt{2}$
	$\approx 13.18 - 0.762n + 0.0822n^2$
$B/P = 1/3(c_{11} + 2c_{12})/P$	$1/3n + 1$
$1/2(c_{11} - c_{12})/P$	$1/8n - 3/4$
	$\approx -0.23 + 0.056n + 0.0024n^2$
c_{44}/P	$1/4n - 1/2$
	$\approx -0.85 + 0.296n - 0.0016n^2$
G_f/P	$-3/5 + 1/5n$
G_{fr}/P	$53/280n - 669/980 - 120/(343n - 1274)$
	$\approx -0.53 + 0.161n + 0.0011n^2$

Derivatives of elastic moduli with respect to pressure for the FCC lattice with central potential $U = A/r^n$ (in this case, $B' = B/P$ etc.). The shear moduli were calculated in the Voigt approximation $G_f = [9c_{44} + 3(c_{11} - c_{12})]/15$ and the Voigt–Royce–Hill approximation $G_{fr} = (G_f + G_r)/2$, where $15/G_r = 12/(c_{11} - c_{12}) + 9/c_{44}$ [4, 5]. The first value in the table is for the short-range (nearest-neighbor) interaction model and the second value is the interpolation in the interval $5 \le n \le 15$ for the usual long-range model. The expressions for B' and G'_f are the same in both cases and are exact. The pressure is normalized to U_a/d^3, where d is the nearest-neighbor separation, and $U_a = A/d^n$ is the nearest-neighbor interaction energy.
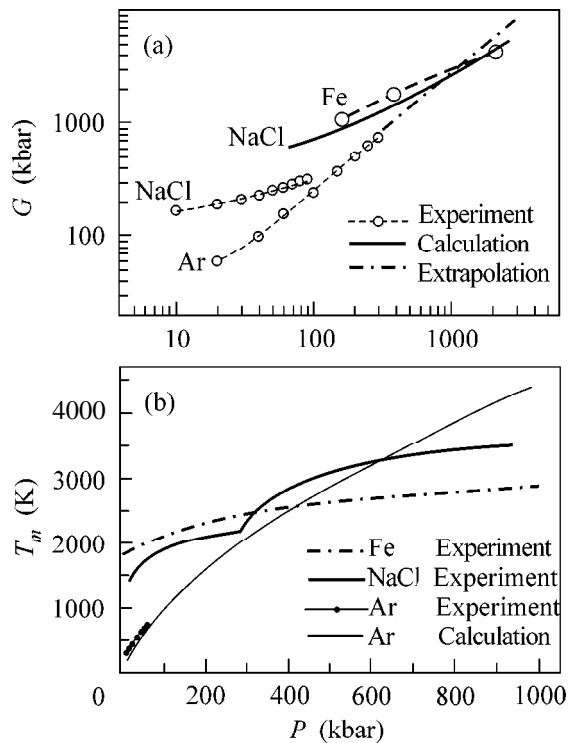
Fig. 3. Pressure dependences of the (a) shear moduli and (b) melting temperatures for typical representatives of rare gas solids (Ar), alkali-halide compounds (NaCl), and metals (Fe). The experimental shear moduli are taken from [6] for Fe or calculated from the experimental elastic constants for Ar [12] and NaCl [14]. The shear modulus of NaCl (with the CsCl structure) is estimated from the theoretically calculated $(c_{11} - c_{12})/2$ values [2]. The theoretical melting curve for Ar is taken from [1], and the experimental curves are taken from [16] for Fe and NaCl (interpolations) and from [21] for Ar.

It follows from the predicted inversion of elastic moduli that, starting at certain pressures, the shear rigidity (hardness etc.) of rare gas solids (except maybe, He) should be higher than the same characteristics of metals and alkali-halide crystals (Fig. 3). Besides, the σ_y/G ratio for metals should be lower because of their high plasticity. Indeed, it is known from experimental studies that Ar becomes stiffer than iron and steel at pressures above 1–1.5 Mbar [19], and Xe becomes stiffer than CsCl at pressures of hundreds of kilobars [20]. The solidified organic carbon-containing liquids polymerize at 50–400 kbar to form diamondlike sp^3 bonds. This should lead to a rapid increase in the shear moduli and hardness up to a level close to diamond. As a result, starting at several hundreds of kilobars, it makes no sense to use organic liquids as a quasi-hydrostatic medium, because they are stiffer at these pressures than practically all materials under investigation.

Thus, the materials that are stiff at normal pressure (metals and ionic dielectrics) are preferable to use as soft quasi-hydrostatic media in the megabar pressure range. The transparent materials such as alkali-halide crystals, and hydrogen-containing compounds such as ice H_2O, ammonia NH_3, etc. are suitable for the optical studies. Note that, although experimental data on the shear modulus of ice H_2O at $P > 100$ kbar are lacking, the melting-curve data for H_2O [21] allows one to assume that ice, along with LiH, becomes one of the softest materials (probably, softer than helium) at $P \sim$ 2–3 Mbar. Indeed, the ionic crystals KBr and NaCl have found use in recent years as pressure-transmitting media in megabar experiments. In many cases, it is reasonable to use the appropriate metals, including In, Pb, Sn, Na, Be, Bi, Ga, etc. (Li is unsuitable because of its active diffusion into the diamond anvils), as quasi-hydrostatic mediums in studying the magnetic, superconducting, and structural properties in the megabar range.

6. In summary, one can assume that, apart from the equalization of elastic characteristics of the materials of various classes, a new phenomenon should occur in the megabar pressure range—inversion of shear rigidity. The molecular substances such as rare gas solids may surpass metals, as well as the ionic and covalent materials, in shear modulus and hardness. Such relative behavior of the moduli should be taken into account when analyzing various processes in the interior of the Earth and planets and also in high-pressure megabarrange experimental technique. Clearly, the predicted phenomenon has a transient character. After the metallization of ionic and molecular materials at pressures of 1–10^3 Mbar, the elastic moduli of all substances will be determined only by the corresponding atomic numbers.

REFERENCES

1. D. A. Young, *Phase Diagrams of Elements* (Univ. of California Press, Berkeley, 1991).

2. V. N. Zharkov and V. A. Kalinin, *Equations of Solid State at High Pressures and Temperatures* (Nauka, Moscow, 1968; Consultants Bureau, New York, 1971).

3. D. M. Teter, MRS Bull., No. 1, 22 (1998).

4. I. N. Frantsevich, F. F. Voronov, and S. A. Bokuta, *Elastic Constants and Elastic Moduli of Metals and Insulators: a Handbook* (Naukova Dumka, Kiev, 1982).

5. M. W. Guinan and D. J. Steinberg, J. Phys. Chem. Solids **35**, 1501 (1974).

6. H.-K. Mao, J. Shu, G. Shen, *et al.*, Nature **396**, 741 (1998).

7. T. S. Duffy, G. Shen, J. Shu, *et al.*, J. Appl. Phys. **86**, 6729 (1999).

8. M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon Press, Oxford, 1954).

9. A. G. Lyapin and V. V. Brazhkin, Phys. Rev. B **54**, 12 036 (1996).

10. V. V. Brazhkin and A. G. Lyapin, Phys. Rev. Lett. **78**, 2493 (1997).

11. C. A. Swenson, in *Rare Gas Solids*, Ed. by M. L. Klein and J. A. Venables (Academic, London, 1977), Vol. II.

12. M. Grimsditch and A. Polian, in *Simple Molecular Systems at Very High Density*, Ed. by A. Polian, P. Loubeyre, and N. Boccara (Plenum, New York, 1989).

13. J. M. Wills and W. A. Harrison, Phys. Rev. B **28**, 4363 (1983).

14. F. F. Voronov, E. V. Chernysheva, and V. A. Goncharova, Fiz. Tverd. Tela (Leningrad) **21**, 100 (1979) [Sov. Phys. Solid State **21**, 59 (1979)].

15. Yu. Ya. Boguslavskiĭ, F. F. Voronov, M. A. Il'ina, and O. V. Stal'gorova, Zh. Éksp. Teor. Fiz. **77**, 946 (1979) [Sov. Phys. JETP **50**, 476 (1979)].

16. R. Boehler, M. Ross, G. Serghiou, and O. Tschauner, Rev. High Pressure Sci. Technol. **7**, 86 (1998).

17. D. Errandonea, R. Boehler, and M. Ross, Phys. Rev. Lett. **85**, 3444 (2000).

18. S. T. Weir, J. Akella, C. Ruddle, *et al.*, Phys. Rev. B **58**, 11 258 (1998).

19. J. Badro, private communications.

20. I. V. Aleksandrov, A. F. Goncharov, I. N. Makarenko, and S. M. Stishov, Zh. Éksp. Teor. Fiz. **97**, 1321 (1990) [Sov. Phys. JETP **70**, 745 (1990)].

21. F. Datchi, P. Loubeyre, and R. LeToullec, Phys. Rev. B **61**, 6535 (2000).

Translated by V. Sakun

Concerted Motion of Protons in Hydrogen Bonds of DNA-Type Molecules[1]

V. L. Golo[1,3], E. I. Kats[2,4], and M. Peyrard[3]

[1] *Department of Mechanics and Mathematics, Moscow State University, Vorob'evy gory, Moscow, 119899 Russia*
e-mail: golo@mech.math.msu.su

[2] *Laue–Langevin Institute, F-38042, Grenoble, France*

[3] *Lab. Phys. ENS Lyon, 69364, Lyon, Cedex 07, France*

[4] *Landau Institute for Theoretical Physics, Russian Academy of Sciences, ul. Kosygina 2, Moscow, 117940 Russia*

Received January 25, 2001

We study the dynamical behavior of proton transfer in hydrogen bonds in the base pairs of double helices of the DNA type. Under the assumption that the elastic and tunnelling degrees of freedom may be coupled, we derive a nonlinear and nonlocal Schrödinger equation (NLNLS) that describes the concerted motion of the proton tunnelling. Rough estimates of the solutions to the NLNLS show an intimate interplay between the concerted tunnelling of protons and the symmetry of the double helix. © *2001 MAIK "Nauka/Interperiodica".*

PACS numbers: 87.14.Gg; 87.15.-v; 33.15.Fm

1. Recent direct observation of coherent proton tunnelling in macromolecules [1] has focused attention on systems allowing for the nontrivial dynamics of protons contained in hydrogen bonds. Perhaps the most significant system of this kind is the DNA molecule. The problem of proton transfer inside the hydrogen bonds of the double helix representing a molecule of DNA has a long and rich history. The importance of the phenomenon was noticed soon after the DNA double helix had been discovered (see [2]). In fact, from the chemical point of view, the proton transfer is a so-called tautomerization reaction that is a kind of transition that preserves the constituent atoms of a compound but, at the same time, changes their mutual positions. It is believed that the tautomerism transition could provide a mechanism for genetic mutations [2].

Later, Crick [3] suggested that the mutations could be due to conformational changes within the double helix, the so-called "wobbling" [2]. By now, there has been no definite conclusion as to whether the wobbling or the tautomerism are responsible for the occurrence of mutations. Recent work on the i-motif of DNA, in which the tautomerism, rare in the usual DNA, is the rule and not an exception [4], has given a new impetus to the problem of proton transfer.

The phenomenon of tautomerism in the base pairs of DNA has been studied extensively, both experimen-

tally and theoretically [2, 5, 6], but for individual single molecules of purines and pyrimidines. Here, it should be noted that even though there is a considerable body of information about the chemical reaction corresponding to the tautomer transition, such as its reaction constant, and even about the concentration of tautomeric forms, 10^{-4}–10^{-5} mol/l, in the B-type of DNA, there is little knowledge of the dynamics of proton transfer accompanying this transition in DNA. In fact, the estimates for the transition frequency widely diverge, and generally it is believed to be within 10^6–10^{11} s^{-1} [4].

In our opinion, the unsatisfactory state of the art, as regards the proton transfer and the tautomeric dynamics in the DNA, is to a certain extent due to the absence of theoretical models of the phenomenon. In this paper, we would like to draw attention to the fact that *if the tautomeric transitions are coupled with a change in elastic properties of the DNA molecule, one could expect a concerted tunnelling of the protons in the hydrogen bonds.* Taking into account the extremely sophisticated nature of the system, i.e., the DNA molecule, we aim at studying it in a qualitative framework allowing for the main features of the molecule, i.e., the presence of *the two strands, the helical structure or the winding symmetry, and the dynamics of protons* considered within the approximation of two-level systems.

2. Aiming at a simple model, we will describe the states of the hydrogen bonds as those of a one-dimensional Bose oscillator, the two-level requirement being accommodated by considering only its ground state and

---

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

the lowest excited one. In this sense, the tautomeric reaction of proton transfer is described by the excited state of an oscillator. Thus, we will consider the protons inside the hydrogen bonds of a molecule of DNA as a quantum system described by the Bose operators b_n^+ and b_n subject to the usual commutation relation

$$[b_n^+, b_m] = \delta_{nm},$$

in which n and m are the indices of the corresponding sites of hydrogen bonds. We will suppose that the system of protons is in a weakly excited state that can be cast in the form of wave function

$$|D(t)\rangle = \sum_n A_n(t) b_n^+ |0\rangle, \tag{1}$$

in which $A_n(t)$ are complex amplitudes satisfying the relation

$$\sum_n |A_n|^2 = 1.$$

The Hamiltonian of the total system, that is the protons and the elastic part corresponding to the sugar phosphates, reads

$$H = H_H + H_Y + H_I, \tag{2}$$

where H_H is the proton Hamiltonian

$$H_H = \sum_n E_0 b_n^+ b_n - \kappa \sum_n (b_{n+1}^+ b_n + b_n^+ b_{n+1}), \tag{3}$$

where E_0 is the level splitting between two states of a proton in a hydrogen bond, κ is the tunnelling probability, H_Y is that of the elastic part, and H_I is the interaction between them. We will consider the elastic part of the system as classical and even neglect its kinetic energy. The reason is that the characteristic times for the elastic vibrations of the DNA molecule are usually estimated as being in the region of 10^{-11}–10^{-13} s, whereas the proton tunnelling is alleged to be within 10^{-6}–10^{-11} s (more exact figures are unavailable; according to [5], the value is 10^{-6} s). Thus, one may suggest that elastic motion follows proton tunnelling without inertia so that we may take into account the lattice deformation only through the potential energy H_Y and neglect the kinetic one. In writing H_Y, we will follow the method worked out in papers [7, 8], aiming at a simplified description of the dynamics of double helix considered as a one-dimensional lattice of vectors $\mathbf{y}_n$ describing mutual positions of the two strands at sites corresponding to the base pair of index n. The helix structure is described with the help of the covariant derivative for the description of deformations resulting from a change

in the positions of the strands, so that the potential H_Y reads

$$H_Y = \sum_n \left[\frac{K}{2}(\nabla \mathbf{y}_n)^2 + \frac{\epsilon}{2}\mathbf{y}_n^2 \right]. \tag{4}$$

Here, the harmonic term $(1/2)\epsilon y_n^2$ describes the binding of the two strands, and $\nabla \mathbf{y}_n$ is the covariant derivative in discrete form, linearized with respect to rotations that relate the two adjacent sites, n and $n + 1$.

$$\nabla \mathbf{y}_n = \frac{1}{a}(\mathbf{y}_{n+1} - \mathbf{y}_n + \hat{\Omega}\mathbf{y}_n) \tag{5}$$

and

$$\Omega_{ij} = -\epsilon_{ijk}\Omega_k, \tag{6}$$

where $\Omega_k \equiv (0, 0, \Omega)$; K and ϵ are elastic constants; and a is the spacing of the lattice of the base pairs. We assume that the molecule is parallel to the z axis. Having chosen the dynamical variables b_n^+, b_n, and $\mathbf{y}_n$ for the protons and the elastic excitations, we have a limited number of options for the interaction energy. The simplest one should couple the proton excited states, given by $b_n^+ b_n$, and the elastic deformations, which are a function of $\mathbf{y}_n$ in the simplest situation we adopt—the linear one. We assume that $\mathbf{y}_n$ enters in the form of its covariant derivative, for this way one might accommodate the interaction of the proton tunnelling with the stack of the π electrons of the bases.

The choice of the interaction Hamiltonian is a delicate problem, but the knowledge of the structure and binding in DNA gives some indications. The stacking interaction between adjacent base pairs is strongly affected by the overlap of the π electrons of the bases. But the proton transfer is accompanied by a redistribution of the electrons on the bases so that the stacking interaction is changed and, therefore, we expect that proton tunneling will affect the *coupling* between adjacent bases. It means that the derivative $\nabla \mathbf{y}_n$, rather than $\mathbf{y}_n$ itself, should enter into the interaction Hamiltonian. We have chosen the simplest form coupling the component of the gradient, which is along the direction $\mathbf{h}_n$ of the hydrogen bonds connecting the bases within the pair n, and the state of the protons in this pair given by $b_n^+ b_n$; i.e.,

$$H_I = -\lambda \sum_n (\nabla \mathbf{y}_n \cdot \mathbf{h}_n) b_n^+ b_n, \tag{7}$$

in which the vectors $\mathbf{h}_n$ are subject to the helical symmetry of the molecule. They can be written in the form

$$\mathbf{h}_n = (\cos n\alpha; \sin n\alpha; 0), \tag{8}$$

where the angle α is related to the helical pitch. Having chosen a primitive model of the DNA, and neglecting subtle features like 2 or 3 hydrogen bonds for the A–T

or G–C base pairs, and confining ourselves to one-site one H-bond picture, we may set $\alpha = \Omega$, equalizing the pitch of $\mathbf{h}_n$ and the twist of the double helix.

It is important to notice that the sign of λ will be crucial for the NLNLS equation that we will derive. Our choice means that the stacking energy is reduced when protons are in the excited state.

We employ the method worked out by Davydov [9] to study the model described above. According to [9], we must calculate the effective potential

$$U_{\text{eff}} = \langle D(t)|H_Y + H_I|D(t)\rangle, \qquad (9)$$

which is a function of $\mathbf{y}_n$; find its minimum $\mathbf{y}_n^0$; and substitute $\mathbf{y}_n^0$ into H given by Eq. (2), thus obtaining the effective Davydov Hamiltonian H_D.

Finally, we must write the Schrödinger equation for the function $|D(t)\rangle$ given by Eq. (1) and the Hamiltonian H_D

$$i\hbar\frac{\partial}{\partial t}\bigg|D(t)\rangle = H_D|D(t)\rangle. \qquad (10)$$

Both sides of the equation indicated above are linear forms in the operators b_n^+, and by equating the coefficients at corresponding b_n^+, we obtain the following equation for A_n:

$$i\hbar\frac{\partial A_n}{\partial t} = E_0 A_n - \kappa(A_{n+1} + A_{n-1})$$

$$- \frac{\lambda^2}{K}|A_n|^2 A_n - \frac{\lambda^2}{K}\left[\sum_m |A_m|^4\right]A_n + \frac{\lambda^2}{2K}\frac{\epsilon a^2}{K\Omega^2}$$

$$\times \left\{\sum_{m',m''} \cos^{|m'-m''|}\phi \right.$$

$$\left. \times \cos[(m'-m'')(\phi-\Omega)]|A_{m'}|^2|A_{m''}|^2\right\}A_n$$

$$+ \frac{\lambda^2}{2K}\frac{\epsilon a^2}{K\Omega^2}\left\{\sum_{m',m''}\left\{\sum_m \cos^{|m-n|}\phi\right.\right.$$

$$\left.\left. \times \cos[(m-n)(\phi-\Omega)]|A_m|^2\right\}A_n\right\},$$

$$(11)$$

in which the angle ϕ is determined by the equation $\tan\phi = \Omega$, and λ/K and $\epsilon a^2/(K\Omega^2)$ are small parameters.

In assessing the importance of cos-terms in the equation given above, it is worthwhile to note that, since we have assumed $\alpha = \Omega$ and there is the relation

$\tan\phi = \Omega$, the typical term in Eq. (11) contains a factor that reads asymptotically

$$\exp(-\Omega^2|m-n|)\cos(|m-n|\Omega^3).$$

We have assumed Ω to be a small parameter and, therefore, the oscillations due to Ω^3 are negligible. Let us consider a possible simplification of the Eq. (11) obtained above. There is a chance that it may have a bearing on the dynamics of the proton transfer in DNA. To this end, note that the angles ϕ and Ω are close to each other; indeed, for the B DNA the angle Ω corresponds to the pitch, i.e., 10 steps for 2π. For this reason, we set all the functions $\cos[(m-n)(\phi-\Omega)]$ equal to 1 in Eq. (11), and introduce

$$A(z,t) = \exp\left(-i\frac{E_0 t}{\hbar}\right)B(z,t). \qquad (12)$$

Then, in the continuous notations, (i.e., for scales larger than a) we obtain the following nonlinear and nonlocal Schrödinger equation ($z \equiv x/a$):

$$i\frac{\partial B}{\partial t} = -\omega_H\frac{\partial^2 B}{\partial z^2} - \frac{\lambda}{K}\omega_T|B^2|B$$

$$- \frac{\lambda}{K}\omega_T B\int|B(z')|^4 dz' + \frac{\lambda}{K}\omega_T\frac{\epsilon a^2}{K\Omega^2}$$

$$\times \int dz' \int dz'' \exp(-\mu|z'-z''|)|B(z')|^2|B(z'')|^2 B(z) \qquad (13)$$

$$+ \frac{\lambda}{K}\omega_T\frac{\epsilon a^2}{K\Omega^2}\int dz' \exp(-\mu|z-z'|)|B(z'|^2)B(z),$$

in which

$$\mu = -\ln\cos\phi = \frac{1}{2}\ln(1+\Omega^2) \simeq \frac{\Omega^2}{2}, \qquad (14)$$

$\omega_H \equiv \kappa/\hbar$, and $\omega_T \equiv \lambda/\hbar$. We may perform a very crude estimate so as to see the part played by the nonlocal terms as regards the structure of solitons that might turn around.

The standard procedure to treat the nonlinear Schrödinger equation is as follows [9, 10]. We are looking for the solution in the propagating wave (soliton) form

$$B = \exp[i(kz-\omega t)]\psi(z-vt). \qquad (15)$$

Introducing Eq. (15) into Eq. (13), we get two equations for the imaginary and real parts of solution (15). From the former, we find the velocity of the soliton v

$$v = 2\omega_H k, \qquad (16)$$

and the real part of (13) leads to the equation

$$\omega\psi = -\omega_H[-k^2\psi + \ddot{\psi}] - \frac{\lambda}{K}\omega_T\psi^3$$

$$-\frac{\lambda}{K}\omega_T\psi\int|\psi|^4 dz' + \frac{\lambda}{K}\omega_T\frac{\epsilon a^2}{K\Omega^2}\psi$$

$$\times \int dz'\int dz''\exp(-\mu|z'-z''|)|\psi(z')|^2|\psi(z'')|^2 \qquad (17)$$

$$+\frac{\lambda}{K}\omega_T\psi\int dz'\exp(-\mu|z-z'|)|\psi(z')|^2.$$

Neglecting μ, we get the standard nonlinear Schrödinger equation, however, with renormalized coefficients. It is easy to find the first integral of the equation, i.e., the energy W

$$(\dot{\psi})^2 + \left[-k^2 + \nu - \frac{\lambda}{K}\xi\left(\frac{\epsilon a^2}{K\Omega^2}C + \Gamma\right)\right]\psi^2$$

$$+\frac{\lambda}{2K}\xi\psi^4 = W, \qquad (18)$$

where we used the following notations:

$$\xi = \frac{\omega_T}{\omega_H}, \quad \nu = \frac{\omega}{\omega_H}, \quad \Gamma = \int dz'|\psi(z')|^4,$$

and C is a factor on the order of 1. It is worth noting that even though we have made rough simplifications as regards the nonlocality, its bearing on the proton dynamics still has remained, as is seen in the $\epsilon a^2/K\Omega^2$ term, preserved in the equation given above. One may infer from the fact that with the approximations used, there is a profound interplay of the dynamics of proton transfer and the conformational structure of the double helix.

From Eq. (17) we obtain the asymptotic width Δ of the soliton, which reads

$$\Delta = \left[\frac{\omega_T^2}{\omega_{ac}\omega_H}\frac{\epsilon a^2}{K\Omega^2}\right]^{-1/2}. \qquad (19)$$

The frequency $\omega_{ac} \equiv K/\hbar$ is generally accepted to be within 10^{11}–10^{13} Hz, but the available estimates for ω_T, ω_H, ϵ widely diverge, $\omega_H = 10^6$–10^{11} Hz, $\omega_T/\omega_H = 0.1$–0.01, and $\epsilon a^2/K\Omega^2 \simeq 0.1$. Consequently, the width of the solution varies within 10–1000 Å, so that one may expect a concerted tunnelling of protons for the lower estimation of Δ.

3. In conclusion, we would like to draw attention to the fact that the existence of an appreciable interaction between the proton transfer inside the hydrogen bonds of the double helix and elastic modes of the latter could result in a concerted dynamics of the protons, which is generally of nonlinear character and governed by the nonlinear and nonlocal Schrödinger equation. To our knowledge, it is the first appearance of the nonlinear Schrödinger equation with nonlocal terms. The transfer of protons may be due to various reasons, and among these are the action of external agents, especially, according to recent experimental work [11, 12], enzymes; it is also worth keeping in mind the possible links with the mutation mechanism [2]. In this paper, we have tried to put these ideas in more quantitative form within the framework of a model that accommodates the basic symmetry structure of the double helix and, as was shown above, allows for certain rough estimates of dynamical features that may surface, presumably, of the soliton nature. In this respect, it is worthwhile to note that even the crude estimate we made conserves the bearing of the double helix, as is seen through the occurrence of the characteristic parameter $\epsilon/(K\Omega^2)$ in the final formulas. Of course, Eq. (11) we obtained has a larger scope. It may not imply the existence of solitons in DNA, but could open another possibility: the existence of nonlinear localization leading to a collapse of an initially broad excitation into a highly localized deformation [13]. Collapse does not occur in the standard NLS equation, but the existence of nonlinearities with higher power in our NLNLS suggests that it could occur in this equation, although, as these terms are nonlocal, no definite conclusion can be given without further studies of the equation. If this hypothesis would be confirmed, a weak and broad perturbation of the hydrogen bonds of a DNA molecule by the vicinity of an enzyme carrying local charges could trigger this nonlinear localization phenomenon and finally lead to the formation of a tutomerized form by the tunneling of one proton. As the present stage of the study, this is however, only a speculation raised by the form of the NLNLS equation that we derived.

REFERENCES

1. A. J. Horsewill, N. H. Jones, and R. Caciuffo, Science **291**, 100 (2001).

2. W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag, New York, 1984).

3. F. H. C. Crick, J. Mol. Biol. **19**, 548 (1966).

4. J. L. Leroy, K. Gehring, A. Kettani, and M. Gueron, Biochemistry **32**, 6019 (1993); S. Bratos, J. Chem. Phys. **63**, 3499 (1975).

5. N. Bodor, M. J. S. Dewar, and A. J. Hagert, J. Am. Chem. Soc. **92**, 2929 (1970).

6. M. Dreyfus, J. Am. Chem. Soc. **98**, 6338 (1976).

7. V. L. Golo and E. I. Kats, Zh. Éksp. Teor. Fiz. **111**, 1830 (1997) [JETP **84**, 1003 (1997)]; V. L. Golo, Yu. M. Evdokimov, and E. I. Kats, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 766 (1999) [JETP Lett. **70**, 780 (1999)].

8. T. Dauxois, M. Peyrard, and A. R. Bishop, Phys. Rev. E **47**, R44 (1993); T. Dauxois, M. Peyrard, and A. R. Bishop, Phys. Rev. E **47**, 684 (1993); T. Dauxois and M. Peyrard, Phys. Rev. E **51**, 4027 (1995).

9. A. S. Davydov, Usp. Fiz. Nauk **138** (4), 603 (1982) [Sov. Phys. Usp. **251**, 898 (1982)].

10. A. C. Scott, Phys. Rev. A **26**, 578 (1982).

11. D. Antoniou and S. Schwarts, Proc. Natl. Acad. Sci. USA **94**, 12360 (1997).

12. A. Kohen and P. J. Klinman, Acc. Chem. Res. **31**, 397 (1998).

13. V. E. Zakharov and V. S. Synakh, Zh. Éksp. Teor. Fiz. **68**, 940 (1975) [Sov. Phys. JETP **41**, 465 (1975)]; J. J. Rasmuseen and K. Rypdal, Phys. Scr. **33**, 481 (1986).