

CONFERENCES AND SYMPOSIA

PACS numbers: 01.10.Fv, **42.25. – p**, **42.30. – d**, 42.60.Jf, 72.80.Vp, **73.20. – r**, **75.30. – m**, **75.50. – y**, 81.05.Bx, 81.05.ue

Modern problems in physical sciences
(Scientific session of the Physical Sciences Division of the Russian Academy of Sciences, 26 October 2011)

DOI: 10.3367/UFNe.0182.201204d.0437

A scientific session of the Physical Sciences Division of the Russian Academy of Sciences (RAS) was held in the conference hall of the P N Lebedev Physical Institute, RAS on 26 October 2011.

The following reports were put on the session's agenda posted on the website www.gpad.ac.ru of the RAS Physical Sciences Division:

(1) **Morozov S V** (Institute of Microelectronics Technology and High Purity Materials, RAS, Chernogolovka, Moscow region) "New effects in graphene with high carrier mobility";

(2) **Volostnikov V G** (Samara Branch of the P N Lebedev Physical Institute, RAS, Samara) "Modern optics of Gaussian beams";

(3) **Mushnikov N V** (Institute of Metal Physics, Ural Branch of the Russian Academy of Sciences, Ekaterinburg) "Intermetallide-based magnetic materials".

The papers written on the base of these reports are published below.

PACS numbers: 72.80.Vp, **73.20. – r**, 81.05.ue
DOI: 10.3367/UFNe.0182.201204e.0437

New effects in graphene with high carrier mobility

S V Morozov

1. Ways of improving mobility and ballistic transport in graphene

The study of graphene is among the most rapidly advancing areas of solid-state physics. Graphene continues to surprise physicists with its various properties, combining a high strength, record high extensibility, and high conductivity, thermal conductivity, optical transparency, etc. But the chief thing attracting the attention of numerous researchers to graphene is its unusual electronic system and unique transport properties [1–5].

Carriers in graphene have exhibited a rather high mobility, even in the early experimental works, despite the fact that a film only one atom in thickness was in no way protected from the environment. It is significant that the high carrier mobility persisted upon increasing the temperature to room temperature owing to relatively weak electron–phonon scattering. However, attempts to raise the mobility to values exceeding the typical values of $(1-2)\times10^4$ cm^2 V^{-1} s^{-1} observed in graphene on oxidized silicon substrates did not meet with success for a long time. The scattering by nanowavy film structure and more recently the scattering by resonant impurities were considered as the main scattering mechanisms responsible for limiting electron mobility in graphene [2, 3].

A new stage in the study of graphene commenced with the emergence of structures exhibiting a far higher (by 1–2 orders of magnitude) mobility. It became clear that the new structures made it possible to study many subtle multiparticle effects, in particular, the fractional quantum Hall effect which was discovered in the new structures [6].

This report reviews the experiments on high-mobility graphene structures, which were recently performed by the author jointly with A Geim's group at the University of Manchester.

Two ways of obtaining more perfect graphene structures began to show: the use of suspended bridges, and encapsulation of graphene between boron nitride (BN) crystallites. The former method involves the making of graphene bridges suspended between metal contacts (Fig. 1a). First, the standard method of micromechanical delamination is employed to fabricate graphene bridges on an oxidized silicon substrate. Then, a part of the oxide under the graphene bridge is removed by chemical etching. An important role is played by the concluding annealing of the structure with the help of current pulses passed through the suspended graphene bridge. The annealing is executed under conditions close to the threshold of bridge damage; however, it permits attaining locally high temperatures (estimated at 700–800 °C) without degrading the remaining parts of the structure. Our experience suggests that long-term exposure of graphene structures on silicon oxide to temperatures above 400 °C impairs the transport properties of graphene.

S V Morozov Institute of Microelectronics Technology and High Purity Materials, Russian Academy of Sciences, Chernogolovka, Moscow region, Russian Federation
E-mail: morozov@iptm.ru

Uspekhi Fizicheskikh Nauk **182** (4) 437 – 455 (2012)
DOI: 10.3367/UFNr.0182.201204d.0437
Translated by E N Ragozin, E G Strel'chenko; edited by A Radzig

Uspekhi Fizicheskikh Nauk **182** (4) 437 – 442 (2012)
DOI: 10.3367/UFNr.0182.201204e.0437
Translated by E N Ragozin; edited by A Radzig

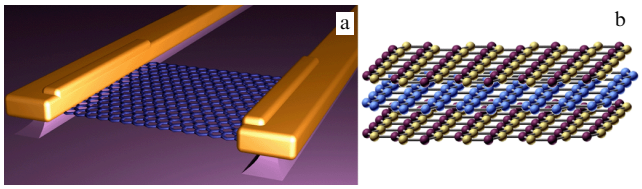

Figure 1. (a) Suspended graphene bridge. (b) Graphene encapsulated between two-dimensional crystals of boron nitride.

This technique permits attaining record high mobility in graphene (above 10^6 cm^2 V^{-1} s^{-1}), but at the same time leads to several disadvantages. First, this technique permits fabricating structures which are highly limited in both the size (typically micrometer-sized) and topology (ordinarily only two-contact structures). Second, suspended graphene structures are in no way protected from the environment as before and the concluding annealing is carried out directly in a low-temperature cryostat in an inert helium atmosphere.

The second way of obtaining high-mobility structures consists in placing graphene between thin (down to monolayer-thin) crystallites of boron nitride (Fig. 1b).

From a more general viewpoint, the latter way of constructing graphene–boron nitride layered structures exhibits a way of fabricating three-dimensional crystals on the base of different two-dimensional constituents. There are different approaches to the formation of new quasi-two-dimensional graphene-based crystals. The feasibility of obtaining new quasi-two-dimensional crystals by chemical modification of graphene has been demonstrated in recent years. This method was employed to obtain graphan [7] and fluorographene [8], in which hydrogen and fluorine atoms, respectively, are attached to the graphene lattice in an ordered manner. Graphan and fluorographene possess dielectric properties, and may be recovered to the initial graphene on thermal processing.

Earlier, we showed [9] that the capacity of forming two-dimensional crystals is inherent not only in carbon and that they may be produced also from other layered materials, like boron nitride, high-temperature superconducting BiSCCO ceramics, and some dichalcogenides. Therefore, the spectrum of two-dimensional crystals is rather broad: from semimetals to dielectrics and superconductors. By placing individual monolayers of different two-dimensional crystals one on top of the other, it is basically possible to obtain new heterostructures with highly diverse properties.

One of the first examples of the implementation of such heterostructures is the encapsulation of graphene between two-dimensional boron nitride crystals with an atomic-smooth surface, which were obtained by micromechanical delamination. The fabrication procedure of the like structures comprises the following stages. First, thin crystallites of boron nitride are placed on the surface of oxidized silicon with the aid of micromechanical delamination. A graphene film is formed in a similar manner on another oxidized silicon substrate preliminarily coated with a polymer. Next, the graphene on the polymer is transferred to the first substrate and superposed with the BN crystallite. On dissolving the polymer, forming the topology of the structure, and metallization routing, the structure is covered with another boron nitride film.

Both methods described above have advantages and disadvantages of their own. Today, the highest electron mobility in suspended graphene bridge structures exceeds 10^6 cm^2 V^{-1} s^{-1}. At the same time, suspended bridges are seriously limited in size as well as in structure topology, which is due to the complexity of their annealing by current pulses, while the two-point measurement system significantly bounds the scope of possible investigations. The graphene on BN is devoid of these drawbacks, since the encapsulation of graphene protects it to a large extent from the action of the environment, but the magnitudes of mobility attained to date have been somewhat lower [as a rule, no greater than $(1-2)\times10^5$ cm^2 V^{-1} s^{-1}].

The advent of high-mobility graphene structures based on boron nitride made it possible to explicitly demonstrate ballistic transport [10]. The experiment was carried out on Hall bridges in the geometry of bend resistance $R_B = V_{34}/I_{21}$ (see the inset to Fig. 2). The current flows between contacts 2 and 1, while the potential difference is measured across contacts 3 and 4, which may be calculated by Van der Pau's formula for the case of diffusion transport. The bend resistance in high-mobility structures is negative in a wide carrier density range, with the exception of a low-concentration domain, testifying that the injected carriers pass from contact 2 to contact 4 almost without scattering (ballistically). A similar effect has earlier been demonstrated in high-mobility GaAlAs heterostructures, and its realization requires that the carrier free path far exceed the characteristic dimensions of the active region of the structure [11]. On application of a weak magnetic field perpendicular to the graphene layer, the bend resistance R_B becomes positive (see Fig. 2), because the magnetic field bends the trajectories of injected electrons and they can no longer reach the opposite contact ballistically. With the assumption of diffuse carrier scattering from the specimen's boundaries, it is possible to estimate the free path in the film bulk: it is ~ 3 μm at helium temperature, and ~ 1.5 μm at room temperature. Therefore, graphene–BN heterostructures exhibit micrometer-scale ballistic transport even at room temperature, which undoubtedly opens up new vistas for the development of high-frequency, low-noise microelectronic applications.

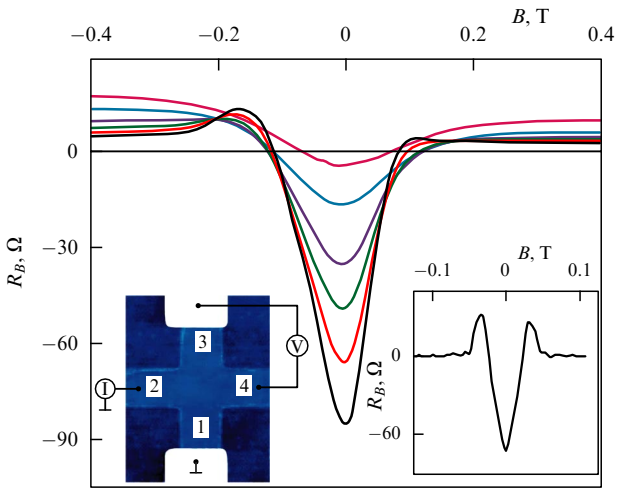

Figure 2. Bend resistance $R_B = V_{34}/I_{21}$ as a function of magnetic field B for $n \approx 6\times10^{11}$ cm^{-2} and temperatures of 50, 80, 110, 140, 200, and 250 K (from the lowermost curve to the topmost one, respectively). The inset at the left shows the measurement schematic, and the inset at the right depicts the calculated dependence $R_B(B)$ for $n \approx 6\times10^{11}$ cm^{-2} in the billiards model [11].

2. Features of the band structure in the vicinity of the electroneutrality point

A high carrier mobility is one of the principal special features of new specimens, though far from being the only one. A wealth of interesting physical effects are expected near the point of electric neutrality (on approaching a zero carrier concentration). But in a real experimental situation, it is hardly possible to reach this point. Indeed, since there are chaotic potential fluctuations, on lowering the carrier concentration the two-dimensional electron gas breaks up into 'pools' with holes and electrons (at points of local rise and fall in the potential, respectively), and therefore a uniform situation with the zero carrier concentration cannot be reached. Of course, how close it is possible to approach the Dirac point depends on the purity and quality of the specimens. New graphene structures have allowed lowering the working carrier concentration by factors of 10–100 and studying the details of the band structure in the vicinity of the electroneutrality point.

First, let us discuss single-layer graphene. The charge carriers in graphene are similar to relativistic particles with a zero rest mass. The linear dispersion law is described by the expression $E = v_F \hbar k$, where the Fermi velocity v_F plays the part of the speed of light, and k is the wave vector. As the carrier concentration lowers and the electroneutrality point is approached, the screening length increases and the electron–electron interaction plays an increasingly important part. The standard Landau theory of a Fermi liquid, which permits representing a strongly interacting electron liquid as a gas of noninteracting electrons, will no longer be applicable. The inset to Fig. 3 illustrates the form of single-layer graphene spectra obtained in a single-particle approximation using the renormalization group technique [12]. The inclusion of electron–electron interaction results in a lowering of the density of states at low energies and in an increase in v_F, with a logarithmic divergence in concentration near the zero energy.

By measuring the temperature dependences of Shubnikov–de Haas oscillations, it was possible to obtain the dependence of the cyclotron mass on the carrier concentration n [12]. The experimental curve $v_F(n)$ corresponding to the latter dependence is depicted in Fig. 3. In new high-mobility specimens there is a possibility (which did not exist in our earlier work [13]) of varying the carrier concentration within three orders of magnitude. Experimental points show a threefold rise in v_F as the Dirac point is approached, which nicely agrees with the theory. Furthermore, several theoretical papers [14–17] predict that electron–electron interaction at low energies may give rise to new electron phases accompanied by the formation of a forbidden band. However, our experiments revealed no hints at a dielectric behavior of graphene down to energies on the order of 0.1 meV.

An even more interesting picture of the dispersion law at low energies was discovered in two-layer graphene [18]. The single-particle theory [19] predicts that the parabolic dispersion law in two-layer graphene passes into a quasilinear one (the four minicones in the left inset to Fig. 4) for an energy $E < 1$ meV. It was not feasible to ascertain the presence of these special features in standard graphene structures on an oxidized silicon substrate, which had a typical mobility of $\sim 10^4$ cm^2 V^{-1} s^{-1}. However, it was possible to verify the data of theoretical calculations on suspended bridges of two-layer graphene with a mobility of $\sim 10^6$ cm^2 V^{-1} s^{-1}. The behavior of the temperature dependence of resistance and the picture of the motion of Landau levels along energy coordinate when varying the magnetic field (Fig. 4) are indications that the cones with a quasilinear spectrum do exist, but there are only two of them rather than four. It is precisely this picture caused by nematic phase transition with a lowering of rotational symmetry in the electron subsystem of two-layer graphene that is predicted by many-particle theories [20, 21].

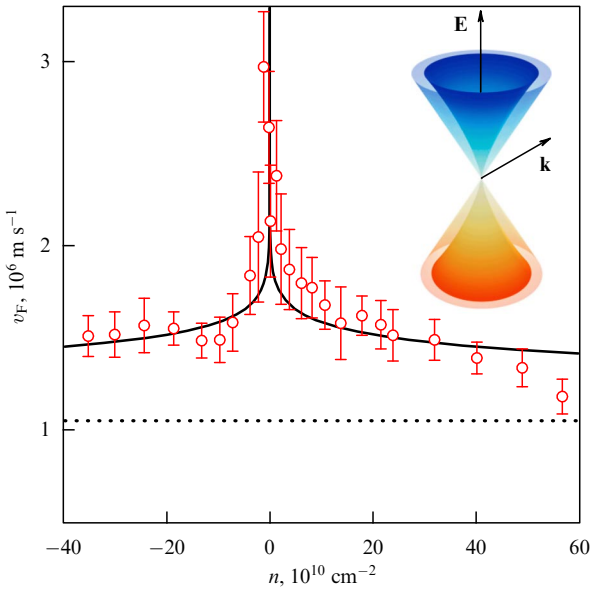

Figure 3. Experimental (circles) and theoretical (solid curve) dependences of Fermi velocity v_F on the charge carrier concentration n in single-layer graphene. The dotted line fits the standard value of v_F in graphene at high concentrations. The inset shows a diagram of the band structure in single-layer graphene at low concentrations without (a single-particle approximation, the outer cone) and with (the inner cone) the inclusion of electron–electron interactions.

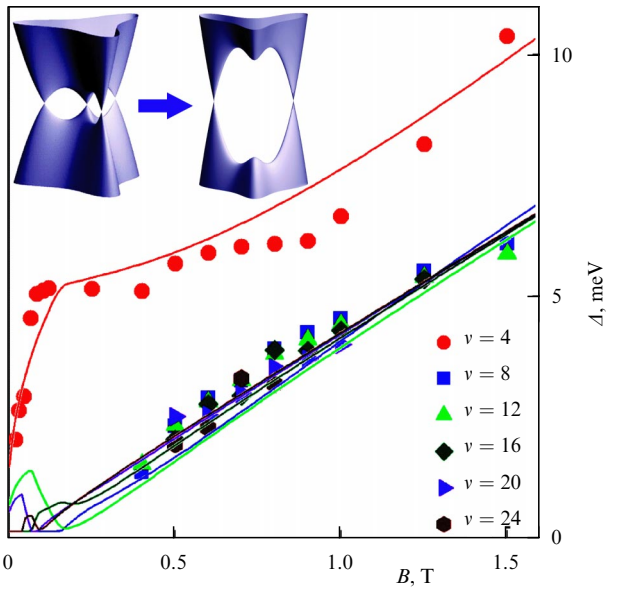

Figure 4. Experimental (symbols) and calculated (curves) dependences of cyclotron gaps Δ for different occupation numbers ν. The inset displays the low-energy portion of the two-layer graphene spectrum in a single-particle approximation (the left part of the inset) and its reconstruction in nematic phase transition (the right part of the inset).

3. Giant spin Hall effect in graphene

The fact that graphene is a semiconductor with a zero band gap revealed itself unexpectedly in nonlocal transport experiments [22]. In the experiment schematized in Fig. 5a, the electric current is passed between the contacts which are quite close to each other, while the potential drop is measured away from the classical current flow path. In a magnetic field, a voltage appeared across remote contacts, whose dependence on the carrier concentration or on the magnetic field was qualitatively similar to the oscillations of longitudinal magnetoresistance. Experiments of this kind had earlier been carried out in high-mobility GaAs heterostructures, and their results were attributed to the edge character of current flow in a quantizing magnetic field and to the scattering of edge states [23, 24]. An advantage of the nonlocal experiment is that it enables filtering out the ohmic contribution in the current flow and detecting more subtle effects.

However, experimental results obtained with graphene significantly depart from those with semiconductors [22]. First, the effect manifests itself with an increase in temperature up to room temperature, and at low temperatures it also takes place in low magnetic fields up to 0.1 T (Fig. 6), making problematic the explanation of experimental data by the scattering of edge states. Second, the effect is much more pronounced only for the zero Landau level (which resides at zero energy).

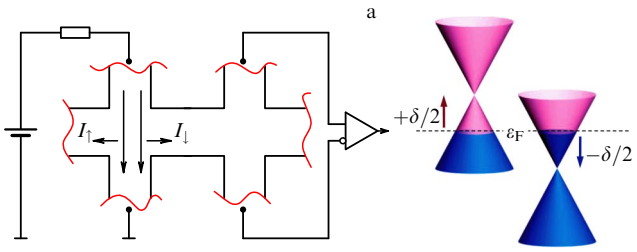

Figure 5. Spin Hall effect in graphene. Schematic of the experiment and an illustration of the emergence of the spin current in a direction perpendicular to the charge current (a) as a result of Zeeman level shifts and the production of electrons and holes with oppositely directed spins in a magnetic field at the electroneutrality point (b).

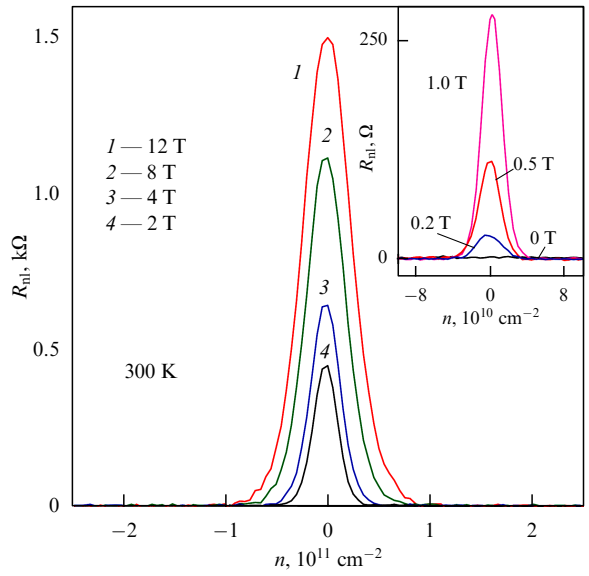

Figure 6. Nonlocal resistance of single-layer graphene encapsulated between boron nitride crystallites in different magnetic fields at room temperature.

The idea of explaining nonlocal response in graphene near the point of electroneutrality is illustrated in Fig. 5b. The Zeeman level splitting in a magnetic field shifts the Dirac cones up and down in energy for quasiparticles with opposite spin directions. When graphene resides at the point of electroneutrality, this gives rise to a finite concentration of electrons and holes with opposite spin directions. In the passage of electric current, the Lorentz force induces the currents of the carriers with opposite spins in the opposite transverse directions, thus leading to a spatial spin disbalance (the spin Hall effect). In this case, the hole and electron charge currents compensate each other in the transverse direction, and the Hall voltage is equal to zero at the point of electroneutrality. Owing to a rather weak spin relaxation [2, 25], a potential drop appears across remote contacts due to inverse transformation.

Phenomenologically, the situation is similar to that for the spin Hall effect caused by spin–orbit interaction [26–28]. The spin–orbit interaction in graphene is weak, and the effect is due to Zeeman level splitting in a magnetic field and the zero band gap. In this case, its magnitude is two orders of magnitude greater than that of the spin Hall effect in conventional semiconductors, which makes the spin Hall effect in graphene quite suited for use in spintronics.

4. Metal–insulator transition in graphene

In the early experimental works on graphene, it came as a surprise that the maximum graphene resistivity ρ near the electroneutrality point approached a value $\propto h/e^2$ (per carrier type), but no indications of strong localization were observed in this case. Indeed, even weak chaotic potential fluctuations in two-dimensional systems lead to a sharp rise in resistivity on lowering the temperature and the carrier concentration [29, 30]. A value of ρ exceeding the quantity h/e^2 signifies that the carrier free path l becomes shorter than the Fermi wavelength λ_F, so that quantum interference begins to play the dominant role, resulting in a strong (Anderson) localization. Because of the thermal activation of the carriers, only a weak temperature dependence of the conductivity was observed, even at the point of electroneutrality in graphene.

The situation changed radically on executing experiments [31] on a system consisting of two closely spaced, but electrically separated, graphene layers with a high mobility ($\sim 10^6$ cm^2 V^{-1} s^{-1}). Figure 7a presents the schematics of the experiment and the multilayer BN–graphene–BN–graphene-BN structure under investigation, which lies on an oxidized silicon plate. Measurements were made of the transport properties of the first (investigated) graphene layer, with the second (control) graphene layer serving as a neighboring gate for the first layer (along with the Si substrate). It is significant that the thickness of the boron nitride layer between the graphene layers was equal approximately to 10 nm, and by varying the carrier concentration in the control graphene layer it was possible to substantially change the screening in the graphene layer under investigation.

At low carrier concentration n_c in the control layer, the graphene layer studied exhibited a usual behavior with a conductivity minimum of $\propto 4e^2/h$. However, for $n_c > 10^{11}$ cm^2 V^{-1} s^{-1} the resistivity rose rapidly near the electroneutrality point at low temperatures (see Fig. 7). In this case, the rise in resistivity was suppressed by a weak

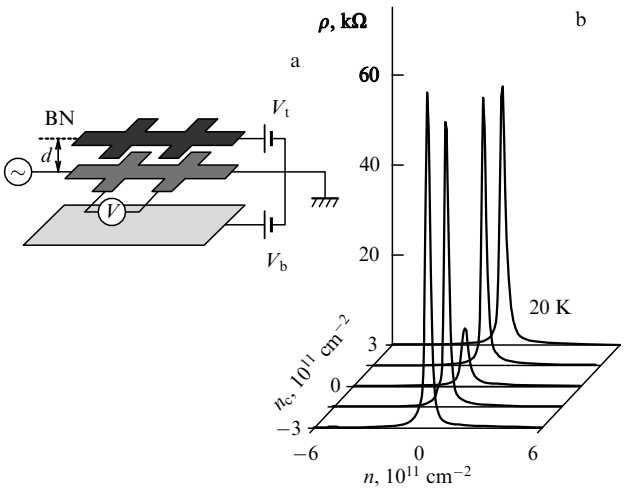

Figure 7. (a) Schematic diagram for the measurement in a structure comprising two separated graphene layers lying on an oxidized silicon substrate (the lower layer). There are boron nitride crystallites (not shown in the drawing) between all the layers. The thickness of the boron nitride layer between the graphene layers is $d \approx 10$ nm. (b) Dependence of the resistivity ρ on the carrier concentration n in the graphene layer under investigation for different carrier concentrations n_c in the control (upper) graphene layer.

magnetic field ($B < 0.1$ T) perpendicular to the graphene layer, which is more likely an indication of an interference effect rather than of the discovery of a forbidden band in graphene.

This behavior of the resistivity is indicative of a metal–insulator transition, demonstrating the Anderson localization on increasing $\rho > h/e^2$. In standard specimens on a silicon substrate, the metal–insulator transition is masked by inhomogeneities and the formation of 'pools' of electrons and holes near the electroneutrality point. The state of graphene inside each of them is far from electroneutrality point, and it remains metallic. Accordingly, the resistivity of the system is determined by semitransparent (owing to Klein tunneling [32, 33]) electron–hole transitions with a weak temperature dependence [34, 35].

The control graphene layer may effectively screen the fluctuation potential and suppress the emergence of electron and hole pools, making it possible to study the behavior of graphene in the vicinity of the electroneutrality point. This interpretation also favors the idea that the proximity of the minimal conductivity in traditional structures to the quantity $4e^2/h$ is due to the flow across the boundaries of the pools of electrons and holes. Therefore, an unusual situation, which is extrinsic to conventional metals and semiconductors, is realized in graphene, whereby localization results from a lowering of disordering rather than from its increasing.

In summary, it should be noted that the emergence of high-mobility graphene structures has led not only to a refinement of certain notions of the graphene physical properties, but also to their revision. At the same time, it is hard to overestimate the promise of the recently commenced work on layered structures consisting of two-dimensional crystals of boron nitride and graphene.

Graphene began its history at some point by separating from its three-dimensional progenitor — graphite. It is not unlikely that in the near future we will obtain a variety of new three-dimensional materials custom-made of different two-dimensional crystals and highly diversified in properties.

Acknowledgments. This work was supported by the Russian Foundation for Basic Research and the Programs of the Presidium of the Russian Academy of Sciences which are gratefully acknowledged.

References

1. Geim A K, Novoselov K S *Nature Mater.* **6** 183 (2007)
2. Castro Neto A H et al. *Rev. Mod. Phys.* **81** 109 (2009)
3. Das Sarma S et al. *Rev. Mod. Phys.* **83** 407 (2011)
4. Lozovik Yu E, Merkulova S P, Sokolik A A *Usp. Fiz. Nauk* **178** 757 (2008) [*Phys. Usp.* **51** 727 (2008)]
5. Morozov S V, Novoselov K S, Geim A K *Usp. Fiz. Nauk* **178** 776 (2008) [*Phys. Usp.* **51** 744 (2008)]
6. Dean C R et al. *Nature Phys.* **7** 693 (2011)
7. Elias D C et al. *Science* **323** 610 (2009)
8. Nair R R et al. *Small* **6** 2877 (2010)
9. Novoselov K S et al. *Proc. Natl. Acad. Sci. USA* **102** 10451 (2005)
10. Mayorov A S et al. *Nano Lett.* **11** 2396 (2011)
11. Beenakker C W J, van Houten H *Phys. Rev. Lett.* **63** 1857 (1989)
12. Elias D C et al. *Nature Phys.* **7** 701 (2011)
13. Novoselov K S et al. *Nature* **438** 197 (2005)
14. Bostwick A et al. *Science* **328** 999 (2010)
15. Khveshchenko D V *Phys. Rev. Lett.* **87** 246802 (2001)
16. Gorbar E V et al. *Phys. Rev. B* **66** 045108 (2002)
17. Drut J E, Lähde T A *Phys. Rev. Lett.* **102** 026802 (2009)
18. Mayorov A S et al. *Science* **333** 860 (2011)
19. McCann E, Fal'ko V I *Phys. Rev. Lett.* **96** 086805 (2006)
20. Lemonik Y et al. *Phys. Rev. B* **82** 201408(R) (2010)
21. Vafek O, Yang K *Phys. Rev. B* **81** 041401(R) (2010)
22. Abanin D A et al. *Science* **332** 328 (2011)
23. McEuen P L et al. *Phys. Rev. Lett.* **64** 2062 (1990)
24. Haug R J *Semicond. Sci. Technol.* **8** 131 (1993)
25. Tombros N et al. *Nature* **448** 571 (2007)
26. Sinova J et al. *Phys. Rev. Lett.* **92** 126603 (2004)
27. Kato Y K *Science* **306** 1910 (2004)
28. Wunderlich J et al. *Phys. Rev. Lett.* **94** 047204 (2005)
29. Imada M, Fujimori A, Tokura Y *Rev. Mod. Phys.* **70** 1039 (1998)
30. Evers F, Mirlin A D *Rev. Mod. Phys.* **80** 1355 (2008)
31. Ponomarenko L A et al. *Nature Phys.* **7** 958 (2011)
32. Katsnelson M I, Novoselov K S, Geim A K *Nature Phys.* **2** 620 (2006)
33. Cheianov V V, Fal'ko V I *Phys. Rev. B* **74** 041403(R) (2006)
34. Adam S et al. *Proc. Natl. Acad. Sci. USA* **104** 18392 (2007)
35. Cheianov V V et al. *Phys. Rev. Lett.* **99** 176801 (2007)

PACS numbers: **42.25.–p**, **42.30.–d**, 42.60.Jf
DOI: 10.3367/UFNe.0182.201204f.0442

Modern optics of Gaussian beams

V G Volostnikov

1. Introduction

A coherent light field, like any oscillatory process, is characterized by its amplitude and phase. The methods and means for analyzing light fields from intensity measurements underlie optical instruments, and from the physical standpoint the solution to any optical measurement problem

V G Volostnikov Samara Branch of the P N Lebedev Physical Institute, Russian Academy of Sciences, Samara, Russian Federation
E-mail: coherent@fian.smr.ru

Uspekhi Fizicheskikh Nauk **182** (4) 442 – 450 (2012)
DOI: 10.3367/UFNr.0182.201204f.0442
Translated by E N Ragozin; edited by A Radzig

involves establishing the relation between the energy and structural parameters of optical radiation.

Owing to the special nature of the optical range, it is not the complex amplitude of an optical signal that is amenable to measurement, but only its intensity, which is not the complete characteristic of a light field in the general case. Traditional interferometric techniques provide, in principle, a possibility of indirect phase measurements; however, in several problems it is impossible or hard to realize the interferometric principle for acquiring information about the complex amplitude or phase of the field. This situation takes place in astronomy, X-ray, and adaptive optics. In this connection, the quest for and investigation of field intensity–phase relations, which provide an answer to the question of how many intensity measurements should be made and of what type should they be so as to reconstruct the field itself or its certain characteristics, are relevant. On the other hand, there is a separate area of investigations in which the sought-after field is to be synthesized rather than reconstructed. This applies to the problems of radiation focusing on a domain with given spatial characteristics and to the task of the intracavity formation of a beam with a given output radiation structure. These problems are kindred to that of light field analysis: they all involve gaining information about the field from its energy characteristics. However, it is easy to see that they are significantly different as well: the physical realizability of a field with the intensity under analysis is inherent in the very formulation of the analysis problem, while the question of whether a field with a given intensity exists is one of the central ones in the synthesis problem. It nevertheless turned out that the results of investigations concerned with the analysis of light fields also open up fundamentally new possibilities for their synthesis in different physical situations. Recent years have seen important new findings, which have not been reflected in monographs. The central aim of this report is to fill in this gap.

This report outlines the results of an investigation into the so-called one-dimensional phase problem in optics. An explicit analytical relation was found between the intensity and the phase of a one-dimensional field in the Fresnel zone.

A two-dimensional problem in optics was also considered and shown to be radically different from the one-dimensional problem. It was established that the vector field of light energy flux generally comprises potential and vortical components; an explicit analytical relation was found between the two-dimensional phase and intensity distributions for vortex-free fields in the Fresnel zone; the vortical component was shown to obey a conservation law, specifically: the integral of the projection of the rotor of the light energy flux vector onto the direction of propagation is equal to zero for any plane in the Fresnel zone, and a relation was revealed between the vortical component of the light energy flux vector and wave-front dislocations.

Also obtained were explicit analytical dependences of the phase of a two-dimensional light field on its intensity as functions of certain parameters of the generating optical system.

The behavior of Gaussian beams under astigmatic action was investigated, too. It was theoretically shown that certain astigmatic optical systems accomplish the mutual transformation of Hermite–Gauss (HG) and Laguerre–Gauss (LG) beams. A parametric class of light beams — the generalized Hermite–Laguerre–Gauss (HLG) beams described by a complete system of parameter-dependent orthogonal functions — was discovered and experimentally realized, the known HG and LG beams being their special types. Optical systems that realize the HG-to-LG beam transformation and the results of the corresponding experiments are described below.

The question of the search for light fields that retain their structure in the course of propagation and focusing, correct to scale and rotation, is formulated and solved in the paraxial approximation. A total description of such light fields, which are termed spiral beams, is given, as are their propagation and rotation laws. The linkage between spiral beams and quantum mechanics is considered. Several ways of realizing spiral beams in experiment are suggested.

Methods for synthesizing light fields with a given intensity distribution that are structurally stable during propagation were considered. Proceeding from spiral beam optics, it was possible to obtain light fields whose intensity distribution is of the form resembling an arbitrary plane curve. The properties of spiral beams for closed curves were investigated. Such beams were found to exhibit characteristic quantization properties: first, the intensity distribution undergoes a radical change under similarity transformations of the corresponding curve and has the shape of this curve only for certain values of the similarity factor; second, the area under the beam's curve for the same values of the similarity factor is related to the Gaussian parameter by an integer-valued relation, and in this case the number of phase singularities of a spiral beam inside the curve is also quantized and their number is defined only by the area inside the curve rather than by its shape.

Also outlined are the results of the application of spiral beam optics to the synthesis problem of phase diffraction optical elements intended for the focusing of a light field on a plane curve and on a two-dimensional domain of a given shape. A new iterative method was proposed for the solution of this problem, which involves the employment of the near-field phase distribution of a spiral beam and its far-field intensity distribution for curves. Proposed for the focusing on a domain were the corresponding distributions of Fourier-invariant fields as the initial approximations in the synthesis of the corresponding phase diffraction optical elements. The results of numerical and natural experiments are presented.

2. Reconstruction of a one-dimensional coherent monochromatic field from measured intensities in the Fresnel zone

Let us elucidate the relationship between the intensity and the phase of a light field in the Fresnel zone. The equation which describes the Fresnel transformation has the form

$$F(x)=\sqrt{\frac{k}{2\pi l}}\,\exp\left(-\frac{1}{4}\,\pi\mathrm{i}\right)\int_a^b \exp\left[\frac{\mathrm{i}k}{2l}\,(x-\xi)^2\right]U(\xi)\,\mathrm{d}\xi\,. \tag{1}$$

Most works on the one-dimensional phase problem are concerned with algorithmic, purely numerical methods for reconstructing the object field from intensity measurements. On the other hand, it would be instructive to elucidate the physical aspect of the problem and its association with the mathematical formulation and, in particular, to derive explicit formulas expressing the intensity–phase relation. This formulation of the problem is all the more justified since the use of explicit formulas offers several advantages from the practical point of view: it shortens the computation

time and permits, in principle, estimating the effect of intensity measurement uncertainty on the accuracy of phase reconstruction.

A similarly formulated problem for a parabolic approximation was studied in several papers (see Ref. [1] and references cited therein), which also suggest that deriving the exact solution requires knowledge of the phase derivative $\partial\varphi/\partial x$ at some point. It is stated simultaneously that finding the boundary condition for $\partial\varphi/\partial x$ from intensity measurements is unlikely.

We show below that it is possible to overcome this difficulty and to determine the field $F(x,l)$ and, hence, the object field $U(\xi)$ from the measurements of intensity $I(x,l)$ and its derivative $I_l(x,l)$ along the direction of field propagation in the Fresnel zone [1]. In the latter zone, $U(\xi)=F(x,0)$ and $F(x,l)$ are related by Eqn (1). It may be shown that the field $F(x,l)$ entering Eqn (1) obeys the parabolic equation [2]

$$\frac{\partial^2 F}{\partial x^2}+2ik\,\frac{\partial F}{\partial l}=0\,. \tag{2}$$

Substituting $F(x,l)$ into Eqn (2) in the form $F(x,l)=\sqrt{I(x,l)}\,\exp\left(i\varphi(x,l)\right)$ and separating the real and imaginary parts give the system of differential equations for the intensity and the phase of the field $F(x,l)$ in the Fresnel zone:

$$\begin{aligned}&\frac{\partial}{\partial x}\left(I\,\frac{\partial\varphi}{\partial x}\right)+k\,\frac{\partial I}{\partial l}=0\,,\\ &2I\,\frac{\partial^2 I}{\partial x^2}-\left(\frac{\partial I}{\partial x}\right)^2-4I^2\left[\left(\frac{\partial\varphi}{\partial x}\right)^2+2k\,\frac{\partial\varphi}{\partial l}\right]=0\,.\end{aligned} \tag{3}$$

The former equation serves as the continuity equation for the flux $\mathbf{j}=(j_x,j_l)=(I\partial\varphi/\partial x,\,kI)$ and expresses the law of light energy conservation in differential form [1]. By integrating this equation, we obtain

$$I(x,l)\,\frac{\partial\varphi(x,l)}{\partial x}=-k\int_{x_0}^{x}I_l(t,l)\,\mathrm{d}t+c\,, \tag{4}$$

where $c=I(x_0)\,\partial\varphi/\partial x\,(x_0)$. Repeated integration yields an expression for the field phase in the Fresnel zone:

$$\varphi(x)=\varphi(a)-k\int_a^x\frac{\mathrm{d}t}{I(t)}\int_{x_0}^{t}I_l(\tau)\,\mathrm{d}\tau+c\int_a^x\frac{\mathrm{d}t}{I(t)}\,. \tag{5}$$

One can see from expression (4) that obtaining the phase requires knowledge of the boundary condition for $\partial\varphi/\partial x$ at some point x_0, the nonlinear character of the relation between $I(x,l)$ and $\varphi(x,l)$ in expression (4) making this issue quite significant. Let us show that the boundary condition for $\partial\varphi/\partial x$ may be found from intensity measurements.

We define the differential operators

$$L\equiv\frac{\partial^2}{\partial z^2}+2ik\,\frac{\partial}{\partial l}\,,\quad L^*\equiv\frac{\partial^2}{\partial z^2}-2ik\,\frac{\partial}{\partial l} \tag{6}$$

and rewrite equation (2) in the form

$$LF(z,l)=L^*F^*(z,l)\equiv 0\,.$$

The action of operators L and L^* on $I(z,l)$ leads to the following result

$$\begin{aligned}&LI(z,l)=2\,\frac{\partial}{\partial z}\left(F(z,l)\,\frac{\partial F^*}{\partial z}\,(z,l)\right),\\ &L^*I(z,l)=2\,\frac{\partial}{\partial z}\left(F^*(z,l)\,\frac{\partial F}{\partial z}\,(z,l)\right).\end{aligned} \tag{7}$$

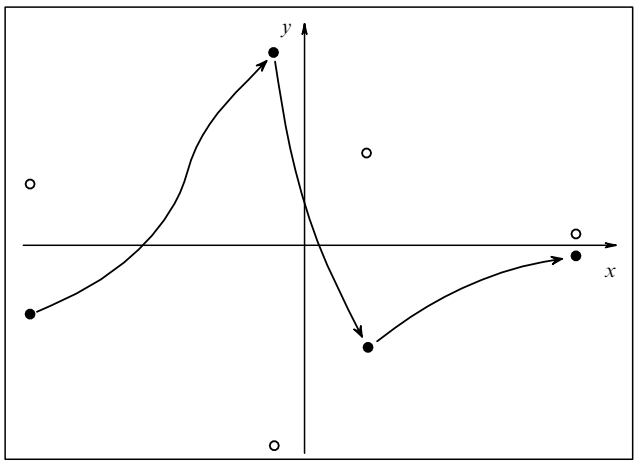

Figure 1. Zero pairs $(z_m,\bar{z}_m)$ of the analytic continuation of intensity $I(z,l)$. Black dots indicate the zeroes of field $F(z,l)$, for which the first of equalities (8) takes place.

Both differential relations (7) are proved by direct substitution with the use of Eqn (2), for instance:

$$\begin{aligned}LI&=F^*LF+FLF^*+2\,\frac{\partial F}{\partial z}\,\frac{\partial F^*}{\partial z}\\ &=2F\,\frac{\partial^2 F^*}{\partial z^2}+2\,\frac{\partial F}{\partial z}\,\frac{\partial F^*}{\partial z}=2\,\frac{\partial}{\partial z}\left(F\,\frac{\partial F^*}{\partial z}\right).\end{aligned}$$

Consequently, if z_1 and z_2 are the zeroes of function $F(z,l)$ for some fixed l, then

$$\int_{z_1}^{z_2}LI(z,l)\,\mathrm{d}z=0\,,\quad\int_{\bar{z}_1^*}^{\bar{z}_2^*}L^*I(z,l)\,\mathrm{d}z=0\,. \tag{8}$$

The use of equalities (8) permits determining all zeroes of the function F from magnitudes of the intensity $I(z,l)$ and its derivative $\partial I(z,l)/\partial l$ for some $l=l_0=\mathrm{const}$ (Fig. 1).

Therefore, the problem of separating out the set of $F(z)$ zeroes from the set of $I(z)$ zeroes may be solved from the distributions of the intensity $I(x)$ and its derivative $I_l(x)$ in some fixed plane $l=\mathrm{const}$ with the aid of analytic continuation and the use of the properties of the functions $I(z)$ and $I_l(z)$ in the complex plane.

3. Reconstruction of a two-dimensional coherent monochromatic field from measured intensities in the Fresnel zone

The linkage between the intensity and the phase of a two-dimensional light field $F(x,y)$ is more poorly understood than its one-dimensional analog. In particular, the nature of the nonuniqueness of the problem's solution and the body of measurements required for its solution are not quite clear.

To substantively analyze the differences between the two- and one-dimensional phase problems, it is expedient to consider the two-dimensional version of a problem solved in the one-dimensional case.

In this section we shall study the two-dimensional problem, or the reconstruction of light field $F(x,y)$ at $l=\mathrm{const}$ from the measurements of intensity $I(x,y)$ and its derivative along the direction of propagation of the radiation $I_l(x,y)$ in the Fresnel zone. This formulation of the phase problem is of interest in the quality control of large-sized

optics, the synthesis of optical elements and light fields, and several other applications.

It is well known that $U(\xi,\eta)$ and $I(x,y,l)$ in the two-dimensional case in the Fresnel domain are related by the expression

$$I(x,y,l) = F(x,y,l)\,\overline{F}(x,y,l)$$
$$= \left| \frac{k}{2\pi i l} \iint_{\Omega} \exp\left(\frac{ik}{2l} \left[(x-\xi)^2 + (y-\eta)^2 \right] \right) U(\xi,\eta)\,d\xi\,d\eta \right|^2 , \quad (9)$$

where Ω is the $U(\xi,\eta)$ carrier, i.e. $U(\xi,\eta) = 0$ when $(\xi,\eta) \notin \Omega$.

The amplitude of $F(x,y,l)$ from expression (9) satisfies the following quasioptical parabolic equation [1]

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + 2ik \frac{\partial F}{\partial l} = 0 . \quad (10)$$

By analogy with the one-dimensional case, $F(x,y,l)$ may be represented in the form $F(x,y,l) = \sqrt{I(x,y,l)} \exp(i\varphi(x,y,l))$ and, on substituting it into equation (10), it is possible to separate the real and imaginary parts.

It is easy to verify that

$$\frac{I\nabla\varphi}{k} = \frac{\overline{F}\,\nabla F - F\nabla\overline{F}}{2ik} = \mathbf{j} = (j_x, j_y) , \quad (11)$$

where $\mathbf{j}$ is the vector component of the light energy flux in the (x, y) plane, and obtain the light energy conservation law, or the continuity equation which may be treated as the three-dimensional divergence of the vector $\mathbf{j} = (j_x, j_y, I)$ or as the equation for a flat field:

$$\frac{\partial \rho}{\partial l} + \nabla \mathbf{j} = 0 , \quad (12)$$

where $\rho = I(x,y,l)$, and $\mathbf{j} = (j_x, j_y)$. Substituting $F(x,y,l) = U(x,y,l) + iV(x,y,l)$ into expression (11) gives

$$\mathbf{j} = \frac{U\nabla V - V\nabla U}{k} = \frac{1}{k} \left(U \frac{\partial V}{\partial x} - V \frac{\partial U}{\partial x} ,\ U \frac{\partial V}{\partial y} - V \frac{\partial U}{\partial y} \right) . \quad (13)$$

One can see from expression (11) that the phase reconstruction problem in this formulation reduces to the reconstruction problem of the vector field from its rotor and divergence. The rotor of the flux is defined as

$$\mathrm{rot}_0\,\mathbf{j} = \frac{2}{k} \left(\frac{\partial U}{\partial x}\frac{\partial V}{\partial y} - \frac{\partial U}{\partial y}\frac{\partial V}{\partial x} \right) = \frac{1}{k} \left(\frac{\partial I}{\partial x}\frac{\partial \varphi}{\partial y} - \frac{\partial I}{\partial y}\frac{\partial \varphi}{\partial x} \right) . \quad (14)$$

Mention should be made of the fundamental difference between the two-dimensional phase problem and the similar one-dimensional one. The solution to the two-dimensional problem is similar to the solution of the one-dimensional problem only when $\mathrm{rot}_0\,\mathbf{j}(x,y) \equiv 0$. This is true, for instance, for fields $F(x,y,l) = F(g(x,y),l)$, where $g(x,y)$ is a real function. The physical meaning of the difference between the two- and one-dimensional cases may be treated as follows. As revealed by analysis of the similar one-dimensional problem, the structural and phase field properties may be directly and completely represented by its energy characteristics; the intensity distribution and the energy conservation law permitted obtaining explicit formulas for the phase of the field. As follows, for instance, from equation (12), the structural and phase properties of the field in the two-dimensional case are not necessarily representable directly by its energy characteristics.

The complete definition of $\mathrm{rot}_0\,\mathbf{j}(x,y)$ in terms of the field intensities is impossible. It is nevertheless possible to prove several results characterizing the local and global properties of $\mathrm{rot}_0\,\mathbf{j}(x,y)$ [3].

Let $F(x,y) = \sqrt{I(x,y)} \exp(i\varphi(x,y))$ be the Fresnel transformation at $l = \mathrm{const}$ of some function with a finite carrier, with the scalar function $\mathrm{rot}_0\,\mathbf{j}(x,y)$ defined by equality (14). Then one has:

1) if $(x_0,\ y_0)$ is the intensity extremum point and $I(x_0,y_0) \neq 0$, then $\mathrm{rot}_0\,\mathbf{j}(x_0,y_0) = 0$; if $I(x_0,y_0) = 0$, then

$$\left|\mathrm{rot}_0\,\mathbf{j}(x_0,y_0)\right| = \frac{1}{k} \sqrt{\frac{\partial^2 I}{\partial x^2}\frac{\partial^2 I}{\partial y^2} - \left(\frac{\partial^2 I}{\partial x\,\partial y} \right)^2} (x_0,y_0) ; \quad (15)$$

2) if $(x_0,\ y_0)$ is an isolated simple zero of function $F(x,y)$ and L is some contour which does not contain zeroes other than $(x_0,\ y_0)$, then

$$\oint_L \nabla\varphi\,d\mathbf{r} = 2\pi\,\mathrm{sign}\,\mathrm{rot}_0\,\mathbf{j}(x_0,y_0) ; \quad (16)$$

3) the following rotor 'conservation law' takes place:

$$\iint_{R^2} \mathrm{rot}_0\,\mathbf{j}(x,y)\,dx\,dy = 0 . \quad (17)$$

In this section, therefore, the problem of the relation between the intensity and the phase of light field $F(x,y,l)$ was considered. In this case, we revealed a radical difference between the two- and one-dimensional phase problem solutions, which is due to the existence of the rotor of the light energy flux vector. The properties of the scalar function $\mathrm{rot}_0\,\mathbf{j}$, which generates this difference, were investigated.

4. Relation between the phase and intensity of the light field as functions of optical system parameters

In Section 3 we showed that passive measurements of the two-dimensional intensity distribution in the Fresnel zone do not permit deriving in general an explicit analytical relation between $I(x,y,l)$ and $\varphi(x,y,l)$. To state it in different terms, the information about the action of the field propagation operator $L = \partial^2/\partial x^2 + \partial^2/\partial y^2 + 2ik\,\partial/\partial l$ on the intensity does not produce a result similar to that obtained for the operator $L = \partial^2/\partial x^2 + 2ik\,\partial/\partial l$ in the one-dimensional case. Now the question is in order: Are there operators which describe real physical situations and provide the solution of this problem for a two-dimensional light field?

Let us consider the transformation of an optical field by an optical system. To simplify calculations, in this section we put $x = x_1,\ y = x_2$. It is well known that the complex amplitude $F(x_1,x_2)$ in the image plane is a Fourier transform of the amplitude $U_0(\xi_1,\xi_2)$ at the exit pupil of an optical system:

$$F(x_1,x_2) = \iint_{R^2} \exp\left[-i(x_1\xi_1 + x_2\xi_2) \right] U_0(\xi_1,\xi_2)\,d\xi_1\,d\xi_2 , \quad (18)$$

where $U_0(\xi_1,\xi_2) = P(\xi_1,\xi_2) f(\xi_1,\xi_2)$, $P(\xi_1,\xi_2)$ is the complex function of the pupil of the optical system, and $f(\xi_1,\xi_2)$ is the Fourier spectrum of the image.

It is evident that $F(x_1, x_2)$ and $I(x_1, x_2) = F(x_1, x_2) \times \overline{F}(x_1, x_2)$ depend on the parameters of the optical system, characterized by function $P(\xi_1, \xi_2)$.

Consider the problem of $F(x_1, x_2)$ reconstruction from the measurements of intensity $I(x_1, x_2)$ as a function of these parameters. By analogy with the aberration theory [1], we represent $P(\xi_1, \xi_2)$ in the form

$$P(\xi_1, \xi_2) = \chi(\xi_1, \xi_2) \exp\left(-\mathrm{i}W(\xi_1, \xi_2)\right), \tag{19}$$

where $\chi(\xi_1, \xi_2)$ is the characteristic pupil function Ω defined as

$$\chi(\xi_1, \xi_2) = \begin{cases} 1, & (\xi_1, \xi_2) \in \Omega, \\ 0, & (\xi_1, \xi_2) \notin \Omega, \end{cases}$$

$W(\xi_1, \xi_2) = W_{21}\xi_1^2 + W_{22}\xi_2^2$; here, W_{2n} stands for astigmatism.

Now if $W_{2n} = \alpha_{2n}$, the problem reduces to the reconstruction of $F(x_1, x_2)$ from the measurements of intensity $I(x_1, x_2)$ and its derivatives with respect to α_{2n} for $n = 1, 2$.

Upon similar transformations, equation (12) reduces to the following system of equations

$$\frac{\partial I}{\partial \alpha_{2n}} + 2\frac{\partial}{\partial x_n}\left(I\frac{\partial \varphi}{\partial x_n}\right) = 0, \quad n = 1, 2. \tag{20}$$

The system of equations (20) describes an optical system with cylindrical phase mask $\exp(-ik\alpha_{2n}\xi_n^2)$ at the exit pupil of the optical system or an optical system with cylindrical defocusing of the illuminating beam and recording in the far-field radiation zone. We note that the problem in the one-dimensional case is completely similar to that considered in Section 2.

Therefore, the field variation under purposeful action on the field is more informative than its variation under natural propagation. This is supposedly an illustration of the well-known fact that an active experiment gives better results than passive observation.

The requisite astigmatic actions (20) on the light field may also be realized directly by employing specific diffraction elements [4].

Let us select the phase function $W(\xi_1, \xi_2)$ entering expression (19) in the form

$$W(\xi_1, \xi_2) = T_0\left(\frac{T}{2\pi}\alpha\xi_1^2 + \xi_1\right) + T_0\left(\frac{T}{2\pi}\alpha\xi_2^2 + \xi_2\right), \tag{21}$$

where $T_0(x)$ is a T-periodic function of argument x, and α is a parameter. The phase element with the profile (21) may be represented as a Fourier series:

$$\exp\left(\mathrm{i}W(\xi_1, \xi_2)\right) = \sum_{m,n} c_m c_n \exp\left(\mathrm{i}m\alpha\xi_1^2 + \mathrm{i}m\frac{2\pi}{T}\xi_1\right) \times \exp\left(\mathrm{i}n\alpha\xi_2^2 + \mathrm{i}n\frac{2\pi}{T}\xi_2\right), \tag{22}$$

where

$$c_m = \frac{1}{T}\int_0^T \exp\left(-\frac{2\pi}{T}\mathrm{i}mx + \mathrm{i}T_0(x)\right)\mathrm{d}x.$$

It is easy to see that such a diffraction element operates as a system of off-axis astigmatic lenses with principal focal lengths $f_m = \pi/\alpha m\lambda$, $f_n = \pi/\alpha n\lambda$ in the (m, n)-th diffraction order (Fig. 2). The diffraction angles of order (m, n) are

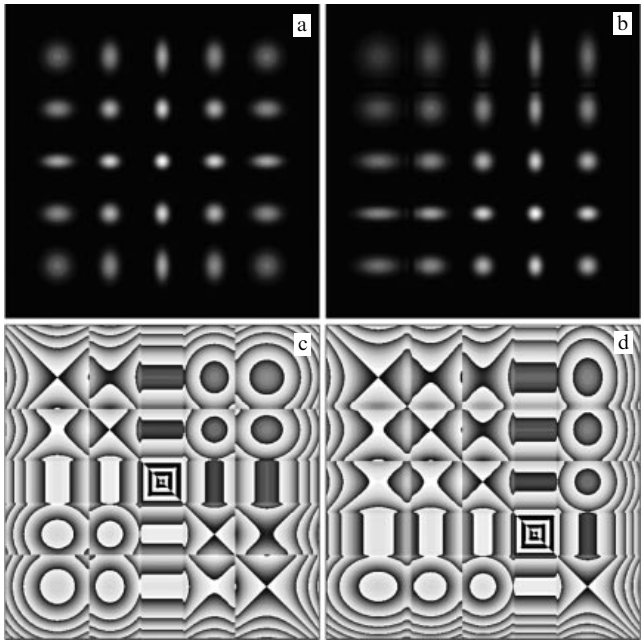

Figure 2. Diffraction pattern of a Gaussian beam: intensity (a, b) and phase (c, d) with diffraction element (21) without astigmatism (a, c), and with astigmatism $\pi(\xi_2^2 - \xi_1^2)/\lambda f_g, f_g = \pi/\alpha\lambda$ in figures (b, d).

$\beta_m = \arcsin(m\lambda/T)$ and $\beta_n = \arcsin(n\lambda/T)$, respectively, and the complex amplitude (19) in the image plane for phase function (21) takes on the form

$$F(x_1, x_2) = \sum_{m,n} c_m c_n \iint_{\mathbf{R}^2} \exp\left(-\mathrm{i}x_1\xi_1 - \mathrm{i}x_2\xi_2\right) U(\xi_1, \xi_2) \times \exp\left(\mathrm{i}m\alpha\xi_1^2 + \mathrm{i}m\frac{2\pi}{T}\xi_1\right)\exp\left(\mathrm{i}n\alpha\xi_2^2 + \mathrm{i}n\frac{2\pi}{T}\xi_2\right)\mathrm{d}\xi_1\,\mathrm{d}\xi_2 = \sum_{m,n} c_m c_n F_{mn}\left(x_1 - \frac{2\pi}{T}m,\ x_2 - \frac{2\pi}{T}n\right). \tag{23}$$

5. Transformation of Hermite–Gauss beams into Laguerre–Gauss beams

The phase problem in optics may be considered as the problem of the linkage between the structural and energy characteristics of a light field. In Sections 2–4 we investigated the relation between the intensity and phase of a light field which satisfies the quasioptical parabolic equation in the Fresnel zone.

Different modifications of this equation describe a broad class of phenomena in quantum mechanics and optics. It is evident that the fields which possess structural stability during propagation occupy a special place, and the intensity–phase relationship is characteristically embodied in them.

On the other hand, as noted in Section 4, an HG beam is transformed into an LG beam in the course of diffraction by an astigmatic diffraction element (Fig. 3). In this connection, the structurally stable solutions of the parabolic equation call for a closer examination.

HG beams are well-known families of the stable solutions of the parabolic equation in optics:

$$\mathcal{H}_{n,m}(x, y) = \exp\left(-x^2 - y^2\right)\mathcal{H}_n(\sqrt{2}x)\,\mathcal{H}_m(\sqrt{2}y), \quad n, m = 0, 1, \ldots, \tag{24}$$

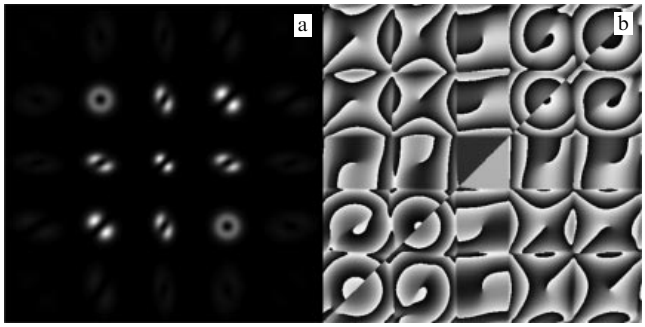

Figure 3. Pattern of Hermite–Gauss beam diffraction by diffraction element (21): intensity (a), and phase (b).

like LG beams:

$$\mathcal{L}_{n,\pm m}(x,y) = \exp(-x^2-y^2)(x \pm \mathrm{i}y)^m \mathcal{L}_n^m(2x^2+2y^2)\,, \quad n,\, m = 0, 1, \ldots, \tag{25}$$

and occupy a prominent place in the theory of resonators and light guides.

Therefore, the change in HG beams under defocusing reduces, correct to a quadratic phase factor, to only a change of scale.

On the other hand, the general form of astigmatism is described by the expression

$$\psi(\xi,\eta,a,\alpha) = a\left[(\xi^2-\eta^2)\,\cos 2\alpha + 2\xi\eta \sin 2\alpha\right]. \tag{26}$$

One can see from expression (26) that, unlike defocusing which is invariant under rotations, the form of the field

$$F(x,y,a) = \iint_{\mathrm{R}^2} \exp\left(-\mathrm{i}(x\xi+y\eta) + \mathrm{i}\psi(\xi,\eta,a,\alpha)\right) \times U(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta \tag{27}$$

depends on the rotation angle α in the propagation in 'astigmatic' space.

Let us now consider the transformation of HG beams under general astigmatic action (see also Ref. [1]):

$$F_{n,m}(x,y,a,\alpha) = \iint_{\mathrm{R}^2} \exp\left(-\mathrm{i}(x\xi+y\eta) + \mathrm{i}\psi(\xi,\eta,a,\alpha)\right) \times \mathcal{H}_{n,m}\left(\frac{\xi}{\rho},\frac{\eta}{\rho}\right)\mathrm{d}\xi\,\mathrm{d}\eta\,. \tag{28}$$

Of particular interest is a special case of transformation (28): for $a = 1/\rho^2$ and $\alpha = \pi/4$, HG beams go over to LG beams [4]:

$$\iint_{\mathrm{R}^2} \exp\left[-\mathrm{i}(x\xi+y\eta) + \frac{2\mathrm{i}\xi\eta}{\rho^2}\right]\mathcal{H}_{n,m}\left(\frac{\xi}{\rho},\frac{\eta}{\rho}\right)\mathrm{d}\xi\,\mathrm{d}\eta = \frac{\pi\rho^2}{\sqrt{2}}\,(-1)^{n+m}\exp\left(-\frac{1}{4}\,\mathrm{i}\rho^2 xy\right) \times \begin{cases} (2\mathrm{i})^n m!\,\mathcal{L}_{m,n-m}\left(\dfrac{\rho x}{2\sqrt{2}},\dfrac{\rho y}{2\sqrt{2}}\right) & \text{for } n \geqslant m\,, \\ (2\mathrm{i})^m n!\,\mathcal{L}_{n,m-n}\left(\dfrac{\rho y}{2\sqrt{2}},\dfrac{\rho x}{2\sqrt{2}}\right) & \text{for } n \leqslant m\,. \end{cases} \tag{29}$$

For any fixed α, the set of fields is a full-value family of orthogonal, structurally stable beams, like the families of HG $\{\mathcal{H}_{n,m}(x,y), n, m = 0, 1, \ldots\}$ and LG $\{\mathcal{L}_{n,m}(x,y), n, \pm m = 0, 1, \ldots\}$ modes. These fields, which were termed generalized Hermite–Laguerre–Gauss beams, were obtained experimentally for different α. These fields were sequentially realized in rotating the cylindrical lens about the optical axis by an angle α. The generalized beams obtained experimentally are exemplified in Fig. 4 for $n = 5$, $m = 4$.

6. Fields with rotation and their properties

As a rule, the alteration of beams in their propagation and focusing is associated with the stretching–compressive deformations: converging and diverging beams. On the other hand, it is evident that even for a simple anisotropy of the beam phase the beam divergence (deformation) also becomes nonuniform. This brings up the legitimate question: Is there some analogy to the torsional strain in the case of a beam with nonuniform divergence? As shown in Section 2, generally the light energy flux consists of two components: divergent, and vortical. In a certain sense, the former component corresponds to stretching–compressive deformations, and the latter one to torsional strains.

In Section 5, we considered the links between HG and LG beams. A characteristic property of these beams is structure retention, correct to scale, in their propagation and focusing. Taking into consideration the vortical component of the light energy flux vector, the notion of structural stability of light fields may be extended. Specifically, this involves examining the question of whether there exist light fields which retain their structure, correct to scale and the character of rotation.

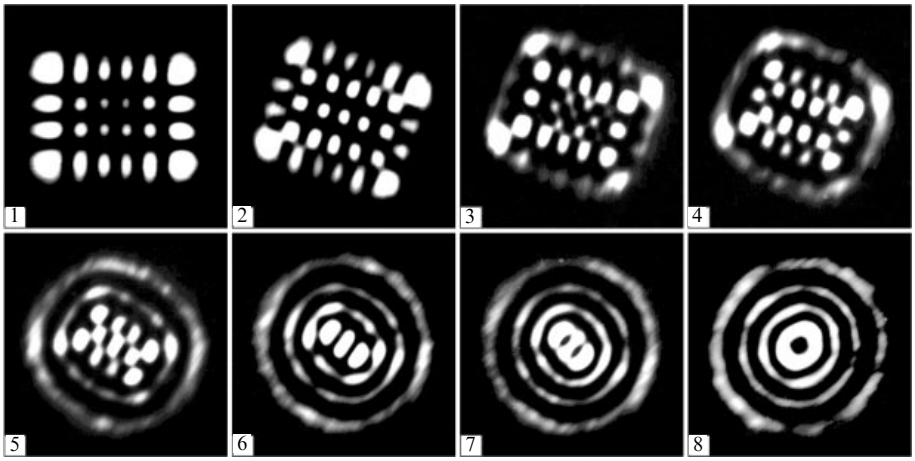

Figure 4. Hermite–Laguerre–Gauss beams with angle α varying from 0 to $\pi/4$.

In this case, the structural stability condition may be defined as follows:

$$I(x,y,l) = D(l) \times I_0\left(\frac{x\cos\theta(l) - y\sin\theta(l)}{d(l)}, \frac{x\sin\theta(l) + y\cos\theta(l)}{d(l)}\right), \quad (30)$$

where $\theta(l)$ stands for the intensity rotation in the propagation of field $F(x,y,l)$, and $d(l) > 0$ is the intensity variation scaling. Let us define real variables by the equality $X + \mathrm{i}Y = (x + \mathrm{i}y)\exp(\mathrm{i}\theta(l))/d(l)$. The exponential decrease in the intensity at infinity (28) permits revealing the structure of the phase $\varphi_0(X, Y, l)$ and rewriting the representation for the light field in the form

$$F(x,y,l) = \frac{1}{d(l)} F_0(X, Y) \times \exp\left(\frac{1}{2}\mathrm{i}kd(l)\,d'(l)(X^2 + Y^2) + \mathrm{i}\gamma(l)\right), \quad (31)$$

where $F_0(X, Y) = \sqrt{I_0(X, Y)}\exp\left(\mathrm{i}\varphi_0(X, Y, 0)\right)$. The structural stability of the intensity thereby generates the structural stability of the phase.

On substituting expression (31) into the parabolic equation, we arrive at the equation for the function $F_0(X, Y)$:

$$\nabla^2 F_0 + 4\mathrm{i}\theta_0\left(X\frac{\partial F_0}{\partial Y} - Y\frac{\partial F_0}{\partial X}\right) - 4F_0(X^2 + Y^2 - \gamma_0) = 0. \quad (32)$$

At $\theta_0 = 0$, equation (32) coincides with the stationary Schrödinger equation for a harmonic oscillator, and its solutions are well known. These are Hermite–Gauss functions $\mathcal{H}_{n,m}(X, Y)$, $\gamma_0 = n + m + 1$ for $n, m = 0, 1, \ldots$, and Laguerre–Gauss functions $\mathcal{L}_{n,m}(X, Y)$, $\gamma_0 = 2n + |m| + 1$ for $n, \pm m = 0, 1, \ldots$.

We shall seek the solutions of equation (32) in the form

$$F_0(X, Y) = \sum_{n,\,\pm m=0}^{\infty} c_{nm}\mathcal{L}_{n,m}(X, Y). \quad (33)$$

Expansion (33) is always possible owing to the finiteness of the energy of field $F_0(X, Y)$ and the completeness of the system of functions $\{\mathcal{L}_{n,m}(X, Y),\ n, \pm m = 0, 1, \ldots\}$ in $L_2(\mathrm{R}^2)$. Substituting expression (33) into equation (32), we obtain

$$\sum_{n,m} c_{nm}\mathcal{L}_{n,m}(X, Y)(2n + |m| + \theta_0 m - \gamma_0 + 1) = 0,$$

and on the strength of the completeness of the system of LG functions, the following relation comes into play:

$$c_{nm}(2n + |m| + \theta_0 m - \gamma_0 + 1) = 0 \quad \text{for all } n, m.$$

Thus, the problem of the search for $F_0(X, Y)$ reduces to the determination of integers n, m from the equation

$$2n + |m| + \theta_0 m = \gamma_0 - 1. \quad (34)$$

Thus, this completes the total description of rotating structurally stable solutions to the parabolic equation, which have come to be known as 'spiral light beams' [2].

The Schrödinger equation for the wave function of a charged particle with mass M and charge e in a uniform magnetic field H has the form

$$\nabla^2\psi + 4\mathrm{i}\,\mathrm{sign}\,(eH)\frac{\partial\psi}{\partial\varphi} - 4\psi\left(R^2 - \frac{2cME_1}{\hbar|eH|}\right) = 0,$$

where $E_1 = E - p_z^2/2M$, E is the particle energy, and p_z is the component of the particle momentum along the field direction. One can see that the last equation is equivalent to Eqn (32).

7. Spiral beams with a given intensity distribution

It is well known from different works on the phase problem that the intensity–phase relations in the one- and two-dimensional cases are radically different. The physical aspects of this difference were considered in Sections 2 and 3, where we came to recognize that it is intimately concerned with the possibility of the occurrence of a vortical component of the light energy flux vector in the two-dimensional approach. A nonzero rotor of the light energy flux vector significantly complicates the intensity–phase relation in this case. On the other hand, this complexity also gives rise to new possibilities.

From the results outlined in Section 6, it follows that in the two-dimensional case there exists a class of coherent light fields — spiral beams — of the form

$$F(x,y,l) = \frac{1}{\sigma}\exp\left(-\frac{x^2 + y^2}{\rho^2\sigma}\right) f\left(\frac{x \pm \mathrm{i}y}{\rho\sigma}\right). \quad (35)$$

One can see from this representation that the class of these fields is rather broad, but the proof of the existence of a beam with given properties and the constructive method of its extraction from this class represents a nontrivial task. This section is concerned with the study of the possibilities for goal-seeking synthesis of light beams (35).

Let us consider several properties of the spiral beams of this class, which follow from representation (35) and are addressed in the subsequent discussion.

Property A. If $\mathcal{S}_n(z, \bar{z}) = \exp(-z\bar{z}/\rho^2) f_n(z/\rho)$ is some totality of spiral beams, their linear combination

$$\mathcal{S}(z, \bar{z}) = \sum_n c_n\mathcal{S}_n(z, \bar{z})$$

is also a spiral beam. And, in general, if $\mathcal{S}_n(z, \bar{z}, a) = \exp(-z\bar{z}/\rho^2) f_n(z/\rho, a)$ stands for a parametric family of spiral beams, then

$$\mathcal{S}(z, \bar{z}) = \int \mathcal{S}(z, \bar{z}, a)\,\mathrm{d}a$$

is also a spiral beam.

Property B. If $\mathcal{S}_0(z, \bar{z}) = \exp(-z\bar{z}/\rho^2) f(z/\rho)$ defines some spiral beam, then

$$\mathcal{S}(z, \bar{z}) = \exp\left(-\frac{z\bar{z}}{\rho^2}\right) f\left(\frac{z\exp(-\mathrm{i}\alpha)}{\rho}\right)$$

likewise is a spiral beam possessing the same intensity distribution as $\mathcal{S}_0(z, \bar{z})$ but turned by an angle α.

Property C. If $\mathcal{S}_0(z, \bar{z}) = \exp(-z\bar{z}/\rho^2) f(z/\rho)$ is some spiral beam, then

$$\mathcal{S}(z, \bar{z}) = \exp\left(-\frac{z\bar{z} - 2z\bar{z}_0 + z_0\bar{z}_0}{\rho^2}\right) f\left(\frac{z - z_0}{\rho}\right) \quad (36)$$

also represents a spiral beam possessing the same intensity distribution as $\mathcal{S}_0(z, \bar{z})$ but shifted to a point z_0. These results naturally generate the following question. Let there be some curve in a plane, specified in a complex parametric form $\zeta = \zeta(t)$, where parameter t varies over the interval $[0, T]$. Does there exist a spiral beam $\mathcal{S}(z, \bar{z}|\zeta(t),\, t \in [0, T])$ with the shape resembling this curve? It turns out that such a beam does exist and assumes the form

$$\mathcal{S}\big(z, \bar{z}|\zeta(t),\, t \in [0, T]\big) = \exp\left(-\frac{z\bar{z}}{\rho^2}\right)\int_0^T \exp\left[-\frac{\zeta(t)\bar{\zeta}(t)}{\rho^2} + \frac{2z\bar{\zeta}(t)}{\rho^2} + \frac{1}{\rho^2}\int_0^t \big(\bar{\zeta}(\tau)\,\zeta'(\tau) - \zeta(\tau)\,\bar{\zeta}'(\tau)\big)\,\mathrm{d}\tau\right]\big|\zeta'(t)\big|\,\mathrm{d}t\,. \quad (37)$$

It is significant that for practical applications, i.e. in the representation of a plane curve in a discrete form, finite increments turn out to be more convenient than differentials, because in the former case there is no need to find derivatives — a none too pleasant task.

The beams for closed curves occupy a special place and deserve separate consideration. Let function $\zeta(t)$, $t \in [0, T]$ describe a closed curve without self-intersections. Without loss of generality, it may be assumed that the curve is traced counterclockwise with increasing t. We define $\zeta(t)$ for all real t by continuing it periodically beyond the segment $[0, T]$. Then, the functions $\zeta(t + a)$, $t \in [0, T]$ describe the same curve for different a. Do the spiral beams for the curves $\zeta(t + a)$ coincide for different a?

It will be shown that the beams constructed for closed curves exhibit characteristic quantization properties. This manifests itself in the fact that, first, the intensity distribution of these curves undergoes radical changes under the similarity transformation $\zeta(t) \to v\zeta(t)$, and possesses the topology of the curve $v\zeta(t)$ only for certain discrete values of parameter v. Second, the intensities of the beams constructed by the curves $v\zeta(t + a)$ for different a are the same only for these values of parameter v.

Let us find the condition whereby the spiral beams constructed for curves $\zeta(t)$ and $\zeta(t + a)$ coincide:

$$\big|\mathcal{S}\big(z, \bar{z}|\zeta(t),\, t \in [a, a + T]\big)\big|^2 \equiv \big|\mathcal{S}\big(z, \bar{z}|\zeta(t),\, t \in [0, T]\big)\big|^2\,.$$

We rewrite the above identity in the form

$$\exp\big(\mathrm{i}\Phi(a)\big)\,\mathcal{S}\big(z, \bar{z}|\zeta(t),\, t \in [a, a + T]\big) \equiv \mathcal{S}\big(z, \bar{z}|\zeta(t),\, t \in [0, T]\big)\,, \quad (38)$$

where $\Phi(a)$ is some real function independent of z (otherwise, by canceling the Gaussian function from both sides of identity (38), we find that Φ is an analytical function of z and, consequently, may not be real for all z). Differentiating identity (38) with respect to a and making use of the periodicity of $\zeta(t)$, one obtains

$$\exp\big(\mathrm{i}\Phi(a)\big)\,\mathcal{S}\big(z, \bar{z}|\zeta(t),\, t \in [a, a + T]\big) \times \left[\mathrm{i}\Phi'(a) - \frac{\bar{\zeta}(a)\,\zeta'(a) - \zeta(a)\,\bar{\zeta}'(a)}{\rho^2}\right] + \exp\left[\mathrm{i}\Phi(a) - \frac{z\bar{z} - 2z\bar{\zeta}(a) + \zeta(a)\,\bar{\zeta}(a)}{\rho^2}\right] \times \left[\exp\left(\frac{1}{\rho^2}\int_0^T (\bar{\zeta}\zeta' - \zeta\bar{\zeta}')\,\mathrm{d}\tau\right) - 1\right]|\zeta'(a)| = 0\,.$$

Replacing the spiral beam in the first term in accordance with identity (38), canceling the Gaussian function, we rewrite this equation in a symbolic form

$$f(z)\,F_1(a) + \exp\left(\frac{2z\bar{\zeta}(a)}{\rho^2}\right)F_2(a) = 0\,,$$

where $f(z)$ is an integer analytical function, and $F_1(a)$ and $F_2(a)$ are some functions of a. This equality takes place for all z, a only when $F_1(a) = F_2(a) \equiv 0$ (when $f(z)$ has a zero, this follows immediately; the case where $f(z)$ has no zeroes is also simple). Therefore, one has

$$\Phi(a) = \frac{1}{\mathrm{i}\rho^2}\int_0^a (\bar{\zeta}\zeta' - \zeta\bar{\zeta}')\,\mathrm{d}\tau,\ \exp\left(\frac{1}{\rho^2}\int_0^T (\bar{\zeta}\zeta' - \zeta\bar{\zeta}')\,\mathrm{d}\tau\right) = 1$$

and, hence [4]

$$\frac{1}{\mathrm{i}\rho^2}\int_0^T \big(\bar{\zeta}(\tau)\,\zeta'(\tau) - \zeta(\tau)\,\bar{\zeta}'(\tau)\big)\,\mathrm{d}\tau = \frac{4S}{\rho^2} = 2\pi N\,,$$

where S is the area bounded by contour $\zeta(t)$.

Therefore, the beam intensity is independent of the origin of integration a only for curves whose area satisfies the quantization condition

$$S = \frac{1}{2}\,\pi\rho^2 N, \quad \text{where} \quad N = 1, 2, \ldots\,. \quad (39)$$

The closed curves which satisfy equality (39) will be referred to as N-quantized curves, and the spiral beams for such curves are called N-quantized beams.[1]

A strictly defined number of optical vortices inside the domain bounded by the generating curve correspond to a quantized beam, which depends on the domain area and not on its shape (Fig. 5) [3, 4]. Hence it follows that, with an increase in domain area, say, from $S = (1/2)\,\pi\rho^2 N$ to $S = (1/2)\,\pi\rho^2(N + 1)$, an increase in the number of zeroes occurs inside the domain due to the entry of one zero from

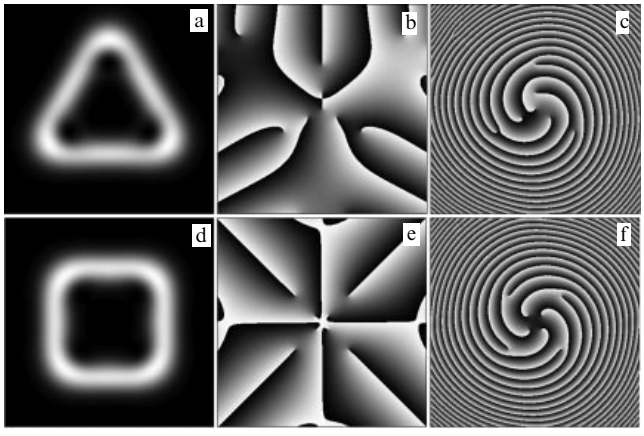

Figure 5. Intensity (a, d), phase (b, e), and phase outside the beam waist (c, f) for a spiral beam in the form of the boundary of a triangle (a–c) and a square (d–f).

[1] If we refer to the quantum-mechanical analogy noted in Section 6, the wave functions of a ground-state particle in a constant magnetic field correspond to the spiral beams with $\theta_0 = \pm 1$, $\gamma_0 = 1$. In this case, condition (39) matches the quantized magnetic flux through the contour $\zeta(t)$: $\Phi = (2\pi\hbar c/|e|)\,N$ (see also Ref. [2]).

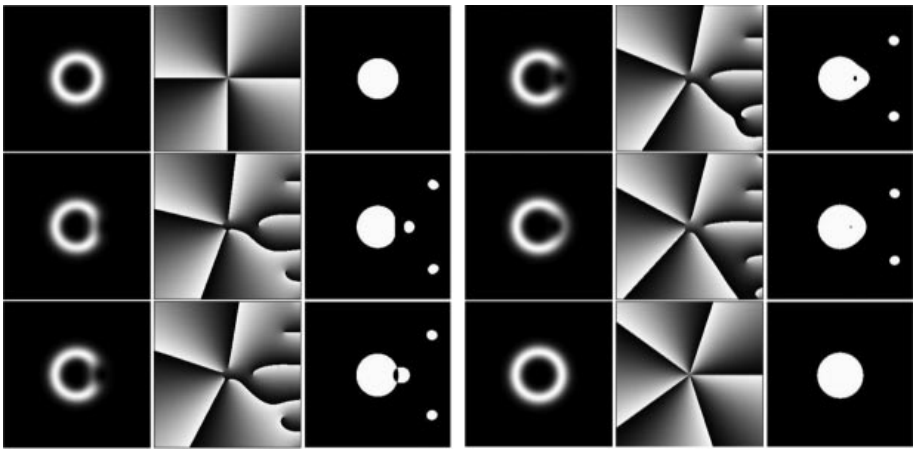

Figure 6. Spiral beam evolution under variation of the radius of the generating circumference.

the outside. One can see the process of zero penetration inside the contour in Fig. 6 which shows the evolution of a spiral beam for a circumference $\zeta(t) = R\exp(\mathrm{i}t)$, $t \in [0, 2\pi]$ for $2R^2/\rho^2 \in [4.0, 5.0]$. The zero entry zone is determined, as discussed above, by the integration origin.

8. Conclusions

The aim of our conclusion consists in providing a summarized and comparative analysis of the new results presented and in formulating some incompletely resolved problems.

The generalized Hermite–Laguerre–Gauss beams found in the investigation of astigmatic transformations of Gaussian beams make up a parametric family in which the previously known HG and LG beams are special representatives corresponding to two certain parameter values. Furthermore, the astigmatic transformations of Gaussian beams permitted proposing a new approach to the synthesis of phase elements intended for the formation of light fields in the form of arbitrarily shaped domains. This problem is now at the research stage, and the level of its solution is still far from the results obtained for light fields in the form of curves. In our view, the reason lies with two interrelated circumstances: first, a domain, unlike a curve, is not an ordered set. Second, light fields in the form of domains contain phase singularities of both signs and are not structurally stable in the Fresnel zone. That is why the synthesis of the appropriate phase elements is complicated by several factors: the domain shape, the synthesis technique, etc.

In recent years, the term singular optics has been used in reference to light fields with wavefront dislocations, or optical vortices. Fields of this kind, which are formed and observed both in linear and nonlinear optical media, are the subject of rather intensive investigation, and the development of adequate theoretical and experimental approaches for the exploration of fields with optical vortices is, therefore, a topical task.

Of course, any coherent light field may be formally represented as a superposition of the known HG and LG beams, but this approach proves to be nonoptimal for the analysis and synthesis of fields with phase singularities.

Vortical light fields which retain, correct to scale and kind of rotation, their structure in the course of propagation, or spiral beams, which are the concern of Section 7, are peculiar 'vortical modes' in the class of fields with phase singularities, and deserve special consideration as a subject of coherent optics.

In our opinion, this is due to the following main reasons. First, spiral beams, despite the fact that they differ greatly in the shape of intensity distributions, are described by explicit analytical expressions, which makes them efficient instruments in the study of the laws of formation and transformation of light fields with phase singularities of a general kind.

Second, in quantum mechanics there is a direct analog to spiral beams — the wave functions of a charged particle in a uniform magnetic field — and the laws of spiral beam transformation have their representation in the theory of coherent states. It is quite possible that these analogies will be mutually beneficial, both for quantum mechanics and for optics. Lastly, the flexibility of variation of spiral beam intensity distributions with the retention of structural stability in the beam propagation and focusing is of interest for laser technologies and the development of specific atomic traps, whereas the nonzero angular momentum of these beams opens new possibilities for manipulating microobjects.

References

1. Volostnikov V G, Preprint No. 93 (Moscow: Lebedev Physics Institute, 1990); *J. Sov. Laser Res.* **11** 601 (1990)
2. Abramochkin E G, Volostnikov V G *Usp. Fiz. Nauk* **174** 1273 (2004) [*Phys. Usp.* **47** 1177 (2004)]
3. Abramochkin E G, Volostnikov V G *Sovremennaya Optika Gaussovykh Puchkov* (Modern Optics of Gaussian Beams) (Moscow: Fizmatlit, 2010)
4. Volostnikov V G, Thesis for the Degree of Doctor of Physicomathematical Sciences (Saratov: Saratov State Univ., 1997)

PACS numbers: **75.30.–m**, **75.50.–y**, 81.05.Bx
DOI: 10.3367/UFNe.0182.201204g.0450

Intermetallide-based magnetic materials

N V Mushnikov

1. Introduction

Intermetallic compounds of transition f- and d-metals are interesting in that they can possess two systems of electrons of different natures, namely localized and itinerant, which can be responsible for magnetism. Rare-earth atoms have their 4f electrons localized both in space and in energy. The interaction of an orbital momentum with the lattice crystal field results in a large magnetic anisotropy and a large magnetostriction. At the same time, the indirect exchange interaction between the 4f electrons is relatively weak, and the magnetic ordering temperatures of rare-earth metals are low. Unlike 4f metals, the energy spectra of transition 3d metals form quasicontinuous bands. 3d (or iron group) metals typically have a high magnetic ordering temperature but a low anisotropy. Rare-earth intermetallic compounds often combine the magnetic characteristics of pure 4f and 3d metals to yield materials with a high ordering temperature and immense anisotropy (or magnetostriction) simultaneously — which enables their use as permanent magnets, as magnetostrictive or magnetoresistive materials, and for magnetic cooling.

The magnetic properties of rare-earth intermetallides have been the subject of intense study since the pioneering work by K P Belov and his team at Moscow State University in the late 1960s. In Ekaterinburg, the initiators of research into the field were S V Vonsovskii, Academician, and Ya S Shur, Corresponding Member of the USSR Academy of Sciences, both at the Institute of Metal Physics (IMP) of the Ural Branch of the USSR Academy of Sciences, and A V Deryagin, Corresponding Member of the USSR Academy of Sciences, at Ural State University. With a host of exciting results obtained in recent years, it is beyond the ambition of this review to cover the entirety of what has been done and achieved in this vast area of the physics of magnetic phenomena. The discussion will be limited to the work carried out at the UrB RAS IMP in close cooperation with other research centers on magnetism of intermetallides, both in Russia and abroad.

2. Permanent-magnet materials

From the middle 1970's onwards, the highest values of specific magnetic energy have been achieved in permanent magnets made of highly anisotropic intermetallide phases such as $SmCo_5$, Sm_2Co_{17}, and $Nd_2Fe_{14}B$ [1]. From the application point of view, of most importance is the segment of high energy-consuming permanent magnets based on Nd–Fe–B [2], whose global production is about 100,000 tons a year. Russia is lagging far behind developed countries both in production volume (about 140 tons a year) and in the property level of rare-earth permanent magnets. In particular, Russian-produced permanent magnets typically have a maximum energy product $(BH)_{\max}$ of 40.0 MG Oe — in contrast to the specific magnetic energy of 50.0 MG Oe accessible on the world market. Recently, the joint research efforts of the IMP, Ural State University, the Ural Electromechanical Plant, and the Moscow Steel and Alloys Institute has resulted in developing a low-oxygen technology for fabricating sintered permanent Nd–Fe–B magnets using strip-cast alloys. With the basic technological processes optimized, a maximum energy product $(BH)_{\max} \geqslant$ 50 MG Oe was obtained industrially, which matches the level of the best magnet manufacturers and is 25% superior to the domestic commercial analogs [3]. Importantly, though, the maximum theoretical value for the $Nd_2Fe_{14}B$ phase is 68.0 MG Oe, leaving significant room for improvement. The maximum theoretical value of the energy product for a permanent magnet is determined by magnetization. The material showing the highest magnetization at absolute zero temperature is holmium, for which the theoretical value of $(BH)_{\max}$ is 350.0 MG Oe. Thus, the potential for magnetic materials based on intermetallic compounds between rare-earth metals and iron group metals is far from exhausted.

High-technology industries including computer hardware, robotics, and automobile engineering are unimaginable without permanent magnets. In the rotors of high-speed centrifuge-driving hysteresis motors, only mechanically strong hard magnetic materials can be used. The IMP Laboratory of Ferromagnetic Alloys developed a high-strength hard magnetic alloy based on nanocrystallized Fe–Cr–Co, whose combination of functional properties makes it unique in the world. Its mechanical properties are close to those of high-strength steels, whereas its magnetic characteristics are optimum for hysteresis motors [4].

In tungsten- and gallium-doped Fe–Cr–Co alloys, by using hardening, annealing, and intense plastic rolling deformation, a multiphase nanocomposite state was formed instead of a conventional modulated structure, resulting in a high-strength high-plasticity material (conditional yield limit of up to 1500 MPa, and relative elongation of up to 6%). Structure studies [5, 6] using nuclear gamma resonance (NGR), X-ray diffraction and transmission electron microscopy established that the hardening of the alloy and the formation of a highly coercive state are due to the decomposition of a homogeneous α-Fe-based solid solution combined with the evolvement of nano-sized particles of the tungsten-enriched paramagnetic intermetallic phase.

3. Materials for magnetothermal applications

The employment of the magnetocaloric effect in cooling machine design [7] is believed in the last few years to be one of the most promising applications of magnetic materials. Unlike the mechanical compressors currently used in refrigerators, magnetic cooling devices are noiseless, consume less electric power, and are ecologically safe. While most of their working prototypes utilize gadolinium as a working medium, a number of rare-earth intermetallides, such as $La(Fe,Si)_{13}H_x$ and RCo_2 (where R is a rare-earth element), also show considerable promise for application. Because the temperature change due to the demagnetization of a material is determined by magnetic entropy, it follows that compounds whose magnetization is strongly temperature-dependent near room temperature hold promise as magnetic cooling materials. This understanding stimulated recent interest of research-

N V Mushnikov Institute of Metal Physics, Ural Branch
of the Russian Academy of Sciences, Ekaterinburg, Russian Federation
E-mail: mushnikov@imp.uran.ru

Uspekhi Fizicheskikh Nauk **182** (4) 450–455 (2012)
DOI: 10.3367/UFNr.0182.201204g.0450
Translated by E G Strel'chenko; edited by A Radzig

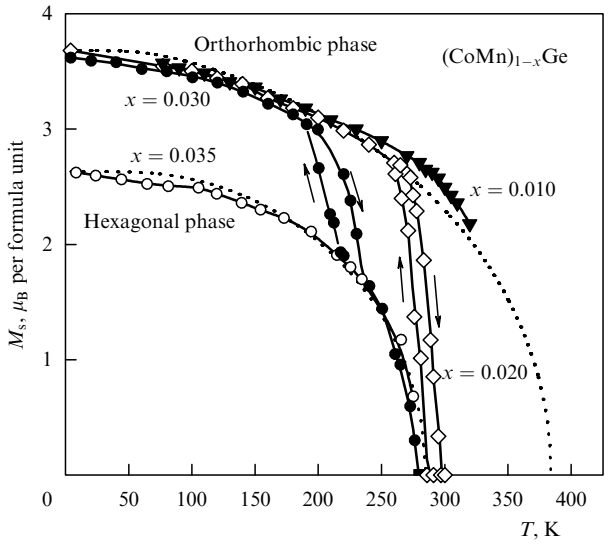

Figure 1. Spontaneous magnetization of $(CoMn)_{1-x}Ge$ compounds as a function of temperature for different x. Dashed lines fit calculated results for the hexagonal and orthorhombic phases.

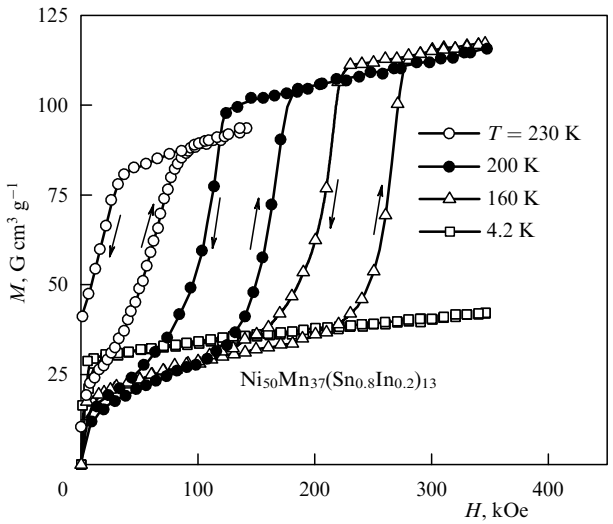

Figure 2. Magnetization curves of the $Ni_{50}Mn_{37}(Sn_{0.8}In_{0.2})$ alloy illustrating a martensite transformation induced by a pulsed magnetic field.

ers in compounds with structural phase transitions near room temperature, with various structural components having widely different magnetizations. The most widely known examples are Heusler Ni_2MnGa alloys [8] and alloys based on $Gd_5(Si,Ge)_4$ [7]. The IMP is currently searching for effective magnetocaloric materials based on CoMnGe, Ni–Mn–Sn–In, etc.

The CoMnGe compound undergoes a diffusionless structural martensite type transition at 470 K from the high-temperature hexagonal Ni_2In type structure to a low-temperature orthorhombic TiNiSi type phase [9]. Although both are ferromagnetic, these modifications differ significantly in magnetization and in the Curie temperature T_C. By varying the composition of the system $(CoMn)_{1-x}Ge$, we were able to obtain alloys for which the structural transition near the ferromagnetic ordering temperature is one from the paramagnetic hexagonal to the ferromagnetic orthorhombic phase [10] (see Fig. 1), which shows exactly the compositions for which magnetothermal effects are expected to be strongest. For $x = 0.02$, the transition occurs in the room temperature range and involves an entropy change of 33.8 J (kg K)$^{-1}$. Applying a magnetic field in excess of the threshold value of 20 kOe in the vicinity of the structural transition increases the relative volume of the higher-magnetization orthorhombic phase.

The magnetic shape memory effect is even more pronounced in Ni–Mn–In and Ni–Mn–Sn alloys [11], in which the austenite and martensite phases differ widely in their magnetization. The $Ni_{50}Mn_{37}(Sn_{0.8}In_{0.2})_{13}$ alloy subjected to an 8-ms magnetic field pulse exhibits a field-induced martensite transformation (Fig. 2) [12]. The critical transition field decreases as the measurement temperature approaches the temperature of spontaneous martensite transformation. Both the spontaneous and the field-induced martensite transformation involve a considerable resistance change: $\Delta R/R \approx -45\%$.

4. Mixed valence compounds and band metamagnets

The situation of practical interest is when two magnetic states of an intermetallide are close in energy, in which case even a slight action on the system (such as an external field, pressure, temperature change, hydrogenation, or doping) can change drastically the magnetic and other physical properties of the system.

A free Yb^{3+} ion has one unfilled 4f electronic state and carries a magnetic moment. When in a compound, the 4f vacancy is often filled by conduction electrons, resulting in the formation of the diamagnetic Yb^{2+} state. In the $YbInCu_4$ compound, the change in valence occurs via a first-order isostructural phase transition in a narrow temperature range [13]. Although the valence of Yb decreases by a mere 0.1, the magnetic susceptibility and electrical resistance show sharp changes, and the specific heat and the sample volume undergo an anomaly (Fig. 3). At low temperatures, applying a magnetic field can destroy the low-magnetization Fermi-liquid state.

Our laboratory study of pressure and doping effects on the transition temperature, transition field, magnetic susceptibility [14], specific heat [15], and magnetostriction [16] shows that the anomaly in physical properties associated with the small change in valence in a given compound is due to a fivefold increase in hybridization factor between f electrons and the conduction electrons [17].

Ferromagnetic ordering in a system of itinerant electrons emerges in the presence of high density of electronic states near the Fermi level. In a band structure of the compound with the Fermi level lying on the steep downward portion of the density-of-states curve, exchange interactions are unstable, and a magnetic field can induce ferromagnetism in a paramagnetic compound — a phenomenon known as band metamagnetism [18]. (This usually requires a strong magnetic field, according to Refs [19, 20].) The YCo_2 compound, being a Pauli paramagnet, changes to a ferromagnetic state in an external magnetic field of 700 kOe [21]. As shown in Ref. [22], doping with aluminium reduces the critical transition field. In Ref. [23], we investigated the effect of hydrostatic pressure on the critical transition field in the $Y(Co, Al)_2$ compounds under pulsed magnetic fields. By comparing the effects of doping, hydrogenation, and pressure, it proved possible to see separately how the interatomic separation and the concentration of d-band electrons contribute to changing the transition field [24].

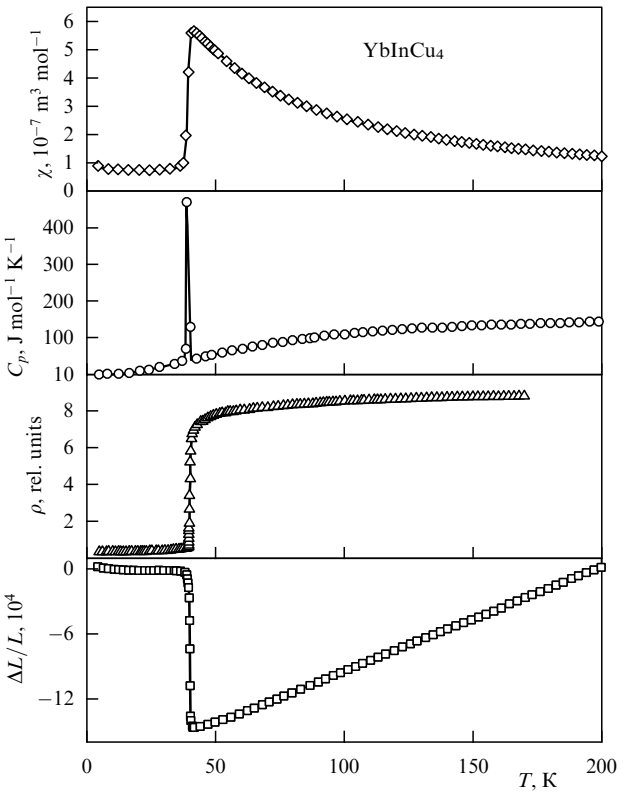

Figure 3. Magnetic susceptibility χ, heat capacity C_p, resistivity ρ, and relative change in the sample size $\Delta L/L$ as functions of temperature for $YbInCu_4$. The isostructural valence-change transition at $T = 40$ K involves a sharp change in properties.

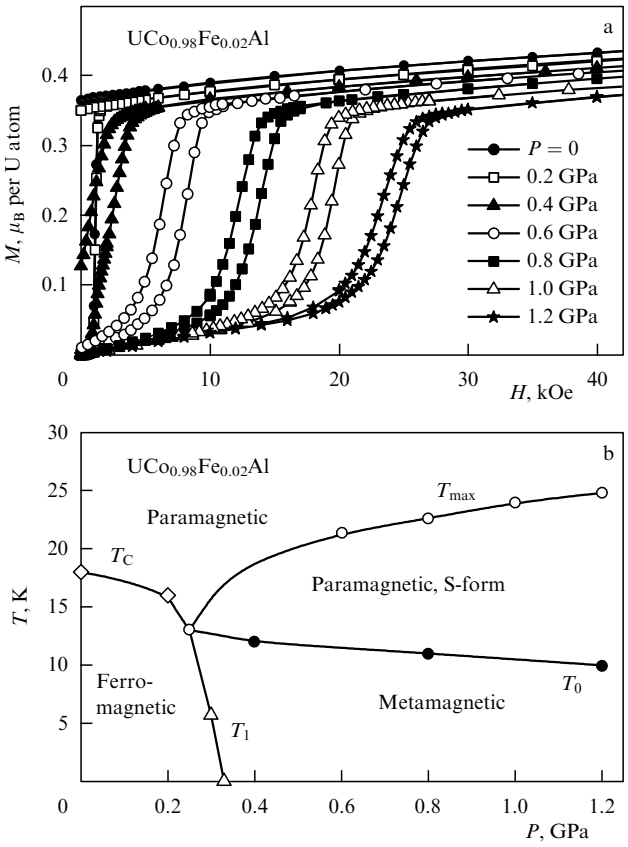

Figure 4. Hysteresis loops at 1.5 K measured for different values of external hydrostatic pressure (a), and magnetic $T-P$ phase diagram (b) for anisotropic band metamagnetic $UCo_{0.98}Fe_{0.02}Al$.

While UCoAl also shows signs of being band metamagnet [25], its properties differ strongly from those of YCo_2. The critical transition field of this compound does not exceed 6.0 kOe. The magnetic moment emerges on uranium, whereas cobalt carries no moment. Metamagnetism is observed only along the c-axis, i.e. is anisotropic, unlike the isotropic metamagnetism seen in YCo_2. The goal of this research, conducted in close cooperation with counterparts from Japan and the Czech Republic, was to reveal the general and particular aspects of an isotropic itinerant 5f-electron metamagnetism.

The metamagnetic magnetization curves of UCoAl measured for different pressures were analyzed with a phenomenological theory of metamagnetic transitions using the magnetic equation of state in the form of Landau expansion in odd powers of magnetization [26]. Matching as best as possible the theoretically predicted to experimental curves yielded the Landau coefficients for different pressures, which determine the temperature behavior of the magnetic susceptibility in the theory of spin fluctuations [27]. The numerical estimates of the maximum magnetic susceptibility and the temperature dependence of the transition critical field are in good agreement with calculated results, implying that spin fluctuation theory is good at describing the properties of anisotropic band metamagnets.

The term 'metamagnet' is also applied to ferromagnets that lose their ferromagnetism — and become metamagnetic — when subject to pressure or acted upon by thermal flux. Such a transition is of particular interest in that applying a relatively weak magnetic field drastically changes the magnetization and conductivity of the material, as well as the size of the sample. In UCoAl, ferromagnetism can be obtained by substituting nonmagnetic iron for only 2% of nonmagnetic cobalt. At a pressure above 0.4 GPa, ferromagnetism is suppressed and, as is typical of UCoAl, a metamagnetic transition occurs (Fig. 4a) and the temperature dependence of the susceptibility reaches a maximum. A study of this sample at various temperatures and pressures provided the full magnetic phase diagram in the pressure–temperature plane [28] (Fig. 4b), which agrees well with the theoretical phase diagram obtained for a spin-fluctuating band metamagnet [29]. To summarize, the results obtained place UCoAl type compounds in a new class of anisotropic itinerant 5f-electron metamagnets.

5. Layered magnets

The study of the physical properties of quasi-two-dimensional and multilayered structures is currently one of the most active areas in the physics of condensed matter. While most research is done on thin multilayered films, there are among intermetallic compounds a number of natural bulk analogs of multilayered structures, including intermetallides of the RM_2X_2 and RM_6X_6 types (R is a rare-earth metal, M is a 3d (4d) transition metal, and X is Si, Sn, or Ge). These compounds consist of intermittent layers of magnetic and nonmagnetic atoms. The strong ferromagnetic interaction of magnetic atoms within a layer results in high magnetic ordering temperatures. Therefore, the magnetic moment of each layer turns out to be a macroscopic quantity and can be treated as a classical vector. At the same time, the relatively weak and easy to control interlayer interaction turns out to be

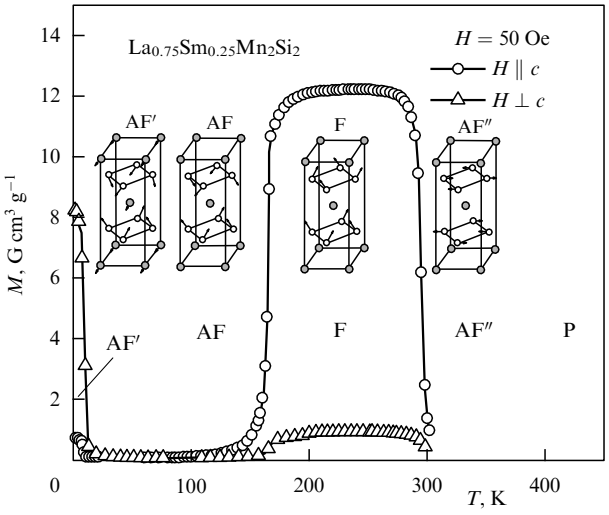

Figure 5. Temperature dependences of magnetization along different single-crystal directions at 50 Oe, and magnetic structures occurring in the $La_{0.75}Sm_{0.25}Mn_2Si_2$ intermetallide at various temperatures. (AF stands for antiferromagnetic structure, F for ferromagnetic, and P for paramagnetic.)

responsible, by and large, for the formation of the magnetic structure of the material.

In RMn_2X_2 intermetallides, the magnitude and sign of the interlayer exchange interaction between the magnetic moments of manganese atoms in neighboring layers depend strongly on the intralayer Mn–Mn distance and hardly, if at all, on the interlayer separation [30]. In the RMn_2Si_2 family, the maximum Mn–Mn interatomic distance is observed for $R = La$, and this is the only compound to exhibit ferromagnetic ordering. Replacing lanthanum by any other rare-earth metal can gradually reduce interatomic distances, thus destabilizing ferromagnetism. The $La_{0.75}Sm_{0.25}Mn_2Si_2$ compound, with the Mn–Mn distance close to the critical value, exhibits a number of spontaneous magnetic phase transitions, due to temperature-related changes in the lattice parameters. In Fig. 5, magnetic structures determined by neutron diffraction are shown for all the phases involved [31]. From the magnetization measurements presented in Fig. 5, one determines the regions of existence of the ferromagnetic and antiferromagnetic phases. In a low-temperature antiferromagnetic state, applying a magnetic field induces a first-order phase transition to a ferromagnetic state [32], producing large changes in the parameters ($\Delta a/a = 1.5 \times 10^{-3}$) and volume of the crystal lattice, as well in the magnetic resistance ($\Delta R/R = -27\%$) [33] (see Fig. 6). Varying the relative content of the alloy components can alter the transition temperatures and critical fields of these compounds over a wide range.

If the interlayer exchange interactions are long-range, then, if there is interaction competition between the nearest neighboring and next-to-nearest layers, a spiral magnetic structure can develop, with the feature (observed in YMn_6Sn_6 [34]) that the spiral divides into two branches shifted through a certain angle. To find out the conditions for the existence of double flat spiral structures, a theory was developed — accounting for three interlayer exchange interactions — which was also employed to model magnetization processes in spiral structures [35, 36]. The results of this modeling provided an explanation for the observed abrupt change in magnetization caused by a magnetic field applied in the basal plane.

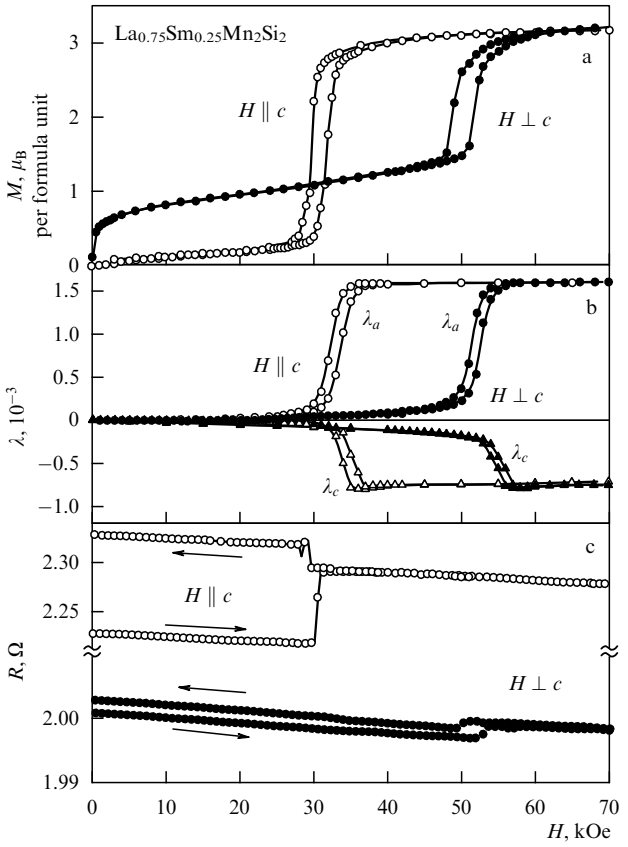

Figure 6. Field dependence of magnetization M (a), linear magnetostriction λ (b), and electrical resistance R (c) along different directions in a single-crystalline $La_{0.75}Sm_{0.25}Mn_2Si_2$ at $T = 4.2$ K.

In a highly anisotropic uniaxial ferromagnet, the application of a magnetic field along the hard magnetization axis normally causes the magnetization vector to turn smoothly in the direction of the field. However, when the magnetic energy of a sample varies nonmonotonically with angle, a field-induced spin-reorientation phase transition involving a sharp change in magnetization occurs in a certain critical field [37, 38]. Unlike most first-order phase transitions, this one exhibits a weak magnetic hysteresis, making it of interest for practical implementations.

The occurrence of magnetic-field-induced first-order phase transitions is thought to be related to a special feature of the compound's magnetic anisotropy: the presence of large higher-order contributions in the expansion of the anisotropy energy in terms of the directional cosines of the magnetization vector. In accordance with the theory of single-ion anisotropy, the higher the order of the contribution, the faster its magnitude decreases with increasing temperature. Therefore, such transitions are usually observed at low temperatures in strong magnetic fields. In the $TbMn_6Sn_6$ compound, the competition among contributions to magnetic anisotropy from the Tb and Mn sublattices, combined with the presence of a spontaneous spin-reorientation transition, result in the occurrence of a magnetic-field-induced first-order phase transition in weak fields around room temperature [39]. This enabled the first neutron diffraction study of how magnetization components change in the transition process [40]. The replacement of terbium by gadolinium in a quasiternary $Tb_{1-x}Gd_xMn_6Sn_6$ system produces a disordered solid solution within the rare-earth sublattice, allowing the magnetic

anisotropy and exchange interactions to be gradually controlled to shift the critical temperature and field of a magnetic phase transition [41].

6. Nonstoichiometric intermetallides

The compounds comprising 4f and 3d metals are usually characterized by a certain stoichiometry and are therefore limited in number. All binary and many ternary intermetallides have been synthesized, and their magnetic characteristics are available in the reference and review literature. At the same time, though, some structures allow deviations from stoichiometry with the result that the properties of a compound can vary widely within a homogeneity region.

It was found recently [42] that the $R\mathrm{Ni_2Mn}$ alloys, where R = Tb, Dy, Ho, and Er, crystallize into an fcc $MgCu_2$ type structure exactly as the compounds $R\mathrm{Ni_2}$ and $R\mathrm{Mn_2}$ do, even though the number ratio of rare-earth-to-transition 3d metal atoms is 1:3. The Mn atoms occupy partly rare-earth and partly Ni sites [43], something which has never been seen before in rare-earth Laves phases. Investigations showed that the $R\mathrm{Ni_2Mn}_x$ alloys remain cubic structure over a wide Mn concentration range [44, 45]. It was found that introducing manganese sharply enhances the exchange interaction and greatly increases the magnetic ordering temperature (4–7 times more than in the original $R\mathrm{Ni_2}$). The maximum T_C is reached for $\mathrm{TbNi_2Mn_{0.75}}$ and equals 160 K.

In the $\mathrm{TbNi_2Mn}_x$ system, spontaneous magnetization monotonically decreases with increasing x. Unlike the original $\mathrm{TbNi_2}$, where the nickel sublattice barely carries a magnetic moment, introducing manganese results in the 3d sublattice acquiring a magnetic moment directed opposite to that of the terbium sublattice. The magnetization curves of ferrimagnetic samples do not saturate until magnetic fields of 15 T. The introduction of manganese results in a monotonic increase in the coercive force which is due to the pinning of narrow domain walls on structural defects. It appears that the partial replacement of terbium atoms by manganese atoms leads to local distortions in the crystal field acting on the Tb ions. As a result, local, random, uniaxial anisotropy arises, giving rise to a noncollinear magnetic structure of the terbium sublattice.

The starting $\mathrm{TbNi_2}$ compound possesses giant magnetostriction at low temperatures [46]. Adding manganese causes the anisotropic magnetostriction to decreases monotonically. Unfortunately, because local anisotropy on Tb sites in these materials is uniaxial, it is only in very strong magnetic fields that magnetostriction due to the rare-earth sublattice can be realized. However, the fact itself of the existence of nonstoichiometric rare-earth intermetallides opens wide possibilities in the search for new magnetic materials.

7. Conclusion

To summarize, the resource of utilizing magnetic materials based on the intermetallic compounds of rare-earth metals with iron group metals is far from exhausted. An important issue is to search for new materials exhibiting large magnetocaloric, magnetostrictive, and magnetoresistive effects. There is still much room for increasing the maximum energy product, extending the operating temperature range, and enhancing the stability of permanent magnets. Today, it is the development of new magnetic materials and of their fabrication technology which determines to a large measure the technological progress.

Acknowledgments. This work was supported by the Russian Foundation for Basic Research (project No. 10-02-96019) and the UrB RAS Program of Fundamental Studies (project No. 12-P-13-2005).

References

1. Deryagin A V *Usp. Fiz. Nauk* **120** 393 (1976) [*Sov. Phys. Usp.* **19** 909 (1976)]
2. Buschow K H J *Rep. Prog. Phys.* **54** 1123 (1991)
3. Popov A G et al. *Fiz. Met. Metalloved.* **109** 257 (2010) [*Phys. Met. Metallogr.* **109** 238 (2010)]
4. Belozerov Ye V et al. *Solid State Phenom.* **152 – 153** 54 (2009)
5. Ivanova G V et al. *J. Alloys Comp.* **509** 1809 (2011)
6. Ivanova G V et al. *Fiz. Met. Metalloved.* **109** 474 (2010) [*Phys. Met. Metallogr.* **109** 438 (2010)]
7. Gschneidner K A (Jr.), Pecharsky V K, Tsokol A O *Rep. Prog. Phys.* **68** 1479 (2005)
8. Khovailo V V et al. *J. Appl. Phys.* **93** 8483 (2003)
9. Johnson V *Inorg. Chem.* **14** 1117 (1975)
10. Markin P E, Mushnikov N V *Solid State Phenom.* **152 – 153** 489 (2009)
11. Kainuma R et al. *Nature* **439** 957 (2006)
12. Gerasimov E G et al. *Solid State Phenom.* **168 – 169** 204 (2011)
13. Felner I et al. *Phys. Rev. B* **35** 6956 (1987)
14. Mushnikov N V et al. *J. Alloys Comp.* **345** 20 (2002)
15. Voloshok T N et al. *Phys. Rev. B* **76** 172408 (2007)
16. Mushnikov N V, Goto T *Phys. Rev. B* **70** 054411 (2004)
17. Mushnikov N V et al. *J. Phys. Condens. Matter* **16** 2395 (2004)
18. Levitin R Z, Markosyan A S *Usp. Fiz. Nauk* **155** 623 (1988) [*Sov. Phys. Usp.* **31** 730 (1988)]
19. Zvezdin A K et al. *Usp. Fiz. Nauk* **168** 1141 (1998) [*Phys. Usp.* **41** 1037 (1998)]
20. Dubenko I S et al. *Pis'ma Zh. Eksp. Teor. Fiz.* **64** 188 (1996) [*JETP Lett.* **64** 202 (1996)]
21. Goto T et al. *J. Appl. Phys.* **76** 6682 (1994)
22. Aleksandryan V V et al. *Zh. Eksp. Teor. Fiz.* **89** 271 (1985) [*JETP* **62** 153 (1985)]
23. Mushnikov N V, Goto T *J. Phys. Condens. Matter* **11** 8095 (1999)
24. Mushnikov N V et al. *J. Magn. Soc. Jpn.* **23** 445 (1999)
25. Sechovský V, Havela L, in *Ferromagnetic Materials: A Handbook on the Properties of Magnetically Ordered Substances* Vol. 4 (Eds E P Wohlfarth, K H J Buschow) (Amsterdam: Elsevier, 1988) p. 309
26. Mushnikov N V et al. *Phys. Rev. B* **59** 6877 (1999)
27. Yamada H *Phys. Rev. B* **47** 11211 (1993)
28. Mushnikov N V et al. *Phys. Rev. B* **66** 064433 (2002)
29. Yamada H, Fukamichi K, Goto T *Phys. Rev. B* **65** 024413 (2002)
30. Szytula A, in *Handbook of Magnetic Materials* Vol. 6 (Ed. K H J Buschow) (Amsterdam: Elsevier, 1991) p. 85
31. Gerasimov E G et al. *Fiz. Met. Metalloved.* **94** (2) 49 (2002) [*Phys. Met. Metallogr.* **94** 161 (2002)]
32. Gerasimov E G, Mushnikov N V, Goto T *Phys. Rev. B* **72** 064446 (2005)
33. Gerasimov E G et al. *J. Phys. Condens. Matter* **20** 445219 (2008)
34. Venturini G, Fruchart D, Malaman B *J. Alloys Comp.* **236** 102 (1996)
35. Rosenfeld E V, Mushnikov N V *Physica B* **403** 1898 (2008)
36. Rozenfel'd E V, Mushnikov N V, Dyakin V V *Fiz. Met. Metalloved.* **107** 599 (2009) [*Phys. Met. Metallogr.* **107** 557 (2009)]
37. Asti G, Bolzoni F *J. Magn. Magn. Mater.* **20** 29 (1980)
38. Belov K P, Zvezdin A K, Kadomtseva A M, Levitin R Z *Orientatsionnye Perekhody v Redkozemel'nykh Magnetikakh* (Orientational Transitions in Rare-Earth Magnets) (Moscow: Nauka, 1979)
39. Terent'ev P B, Mushnikov N V *Fiz. Met. Metalloved.* **100** (6) 50 (2005) [*Phys. Met. Metallogr.* **100** 571 (2005)]
40. Zaikov N K et al. *Pis'ma Zh. Eksp. Teor. Fiz.* **72** 623 (2000) [*JETP Lett.* **72** 436 (2000)]
41. Terent'ev P B et al. *J. Magn. Magn. Mater.* **320** 836 (2008)
42. Wang J L et al. *Phys. Rev. B* **73** 094436 (2006)
43. Mushnikov N V et al. *Phys. Rev. B* **79** 184419 (2009)
44. Mushnikov N V et al. *Fiz. Met. Metalloved.* **110** 223 (2010) [*Phys. Met. Metallogr.* **110** 210 (2010)]
45. Mushnikov N V et al. *Solid State Phenom.* **168 – 169** 200 (2011)
46. Abbundi R, Clark A, McMasters O *IEEE Trans. Magn.* **16** 1074 (1980)