Physics – Uspekhi **39** (6) 635 – 642 (1996)

CONFERENCES AND SYMPOSIA

Scientific session of the Division of General Physics and Astronomy of the Russian Academy of Sciences (31 January 1996)

A scientific session of the Division of General Physics and Astronomy of the Russian Academy of Sciences was held at the P L Kapitsa Institute of Physical Problems on 31 January 1996. Three papers were presented at this session:

(1) **K V Kovalenko, S V Krivokhizha, I L Fabelinskiĭ, L L Chaikov** (P L Lebedev Physical Institute of the Russian Academy of Sciences) "Some features of phase transitions in solutions with two critical points"

(2) **E L Nagaev** (Institute for High-Pressure Physics of the Russian Academy of Sciences), **V V Osipov, A A Samokhvalov** (Institute for Physics of Metals, Ural Branch of RAS) "Cooperative electric phenomena in degenerate magnetic semiconductors with spontaneous phase separation"

(3) **L A Khalfin** (St.-Petersburg Division of the V A Steklov Mathematical Institute of the Russian Academy of Sciences) "The quantum theory of wave packet scattering, the causality principle, and superlight tunnelling"

An abridged version of the papers is given below.

PACS numbers: **64.60.–i**, 64.60.Kw, **64.75.+g**

Some features of phase transitions in solutions with two critical points

K V Kovalenko, S V Krivokhizha, I L Fabelinskiĭ, L L Chaikov

1. Among the great variety of diverse liquid solutions one finds binary solutions with critical mixing points.

If on the temperature versus concentration (TX) phase diagram the coexistence curve for the separated phases and homogeneous solution has a peak, then the solution possesses an upper critical point which is determined by the maximum of the coexistence curve.

If the coexistence curve has a valley, a lower critical point exists, which corresponds to the minimum of the curve. Both classes of solutions have been widely investigated experimentally and theoretically [1 – 4].

There also exist binary solutions, such as guaiacol-glycerol, which are homogeneous over the entire TX phase plane. But adding to such a solution even a minute amount of a certain third component produces a closed region, or loop in its TX diagram, within which the solution is laminated, i.e. heterogeneous, whereas over the rest of the plane it is left homogeneous. It such a solution, two critical points, an upper and a lower, coexist.

Uspekhi Fizicheskikh Nauk **166** (6) 683–690 (1996)
Translated by E G Strel'chenko, edited by A Radzig

2. A theory describing the phase diagrams of solutions such as glycerol-bensylethylamine, glycerol-*m*-toluidine, guaiacol-glycerol, and the like, was developed by American physicists [5]. From the point of view of this theory, the reason why guaiacol and glycerol form a homogeneous solution over the whole of the phase plane is that there are very strong hydrogen bonds acting between the components. Adding to such a solution a small amount of water results in some of the hydrogen bonds of the guaiacol being 'closed' onto those of the water, thus leading to a guaiacol-glycerol lamination. Consequently, molecules of the same species have greater binding energy than molecules of different species.

The existence of two critical points also requires explanation. Such an explanation has been given [5] and lies in the fact that the free energy $F = E - TS$ (where E is the total energy of the system, and S is its entropy) passes a minimum at the critical point.

Moving from the heterogeneous into the homogeneous phase toward the upper critical point increases the temperature as well as gives rise to a small increase in the entropy, with the result that TS becomes large enough to minimise F.

If one moves from the heterogeneous to the homogeneous phase through the lower critical point, T is small and cannot minimise the free energy F at this point. One way out of this difficulty [5] is to introduce a certain orienting force whose role is to reduce the total energy E. It is assumed that this is the hydrogen binding force. Above the upper critical point this force is assumed to be small, whereas below the lower critical point it is so strong that, in fact, the minimization of the total energy E takes place. Our qualitative spectral studies [8] supported this hypothesis.

3. Our previous experimentation with the water-added solution of guaiacol-glycerol (G-G) was concerned with the determination of the critical index for the correlation radius of the concentration fluctuation [7]. The purpose of the present work is the experimental investigation on the properties of the G-G solution. Following the earlier assumption [5] that the water hydrogen bonds may lead to a lamination loop in the G-G solution, then, given that alcohol is also rich in hydrogen bonds, it is hoped that adding the latter instead of water will also form a laminated region.

In actual fact, a quite different picture emerges: the addition of alcohol to the G-G solution does not lead to the formation of a laminated region. Indeed, if a G-G solution already offers a laminated region due to the admixture of water, adding any kind of alcohol leads to the disappearance (collapse) of this region.

If it is the interplay of the hydrogen bonds which has produced the laminated region in the G-G solution, then the addition to the 'dry' solution of a neutral liquid such as carbon tetrachloride (CCl_4) should not affect the properties of the solution. The actual picture is totally different, however. The

addition of carbon tetrachloride to the G-G solution forms a laminated region, whose size, as in the case of water, depends on the precise amount of the third component, CCl_4.

In order to form a laminated region in which the temperature separation between the upper and lower critical points is $\Delta T \approx 7$ K, it is enough to have one molecule of water for 23 solution molecules, whereas in the case of CCl_4, the ratio of 1 to 170 will do. It seems surprising that such small quantities of the third component affect solution properties as strongly as they do. Although the exact mechanism of the processes involved is now difficult to understand, indications are that it must be of the trigger type.

One further important observation is that if the substance of the third component dissolves only in one of the two main solution components (water in glycerol, and carbon tetrachloride in guaiacol), then a laminated region develops in the solution. If, on the other hand, the third component dissolves in both main components (as alcohol, for example), no lamination appears.

The experiments and observations made suggest that either hydrogen bonding has nothing to do with the mechanism of laminated region formation or this mechanism differs from that conceived in Ref. [5].

Experiments involving the addition of CCl_4 demonstrate that hydrogen binding forces cannot account for the existence of the lower critical point. On the other hand, given the absence of chemical reactions in the G-G solution, it is reasonable to assume that at temperatures close to the lower critical point certain structures form in the solution which reduce markedly the total system energy E and are different from the structures existing above the upper critical point.

4. Presumably, some insight into the possible difference in the structure of the homogeneous solution above the upper and below the lower critical points may be gained by measuring the temperature dependence of the velocity of sound, because this latter is determined by the equation of state of the homogeneous solution. In fact, the velocity of sound is, by definition, $V = (\mathrm{d}P/\mathrm{d}\rho)_s^{1/2}$, where P is the pressure, and ρ is the density. The subscript s denotes the adiabatic value of the quantity so labelled.

In our experiments, the hypersonic velocity was derived from the position of the Mandelstam–Brillouin (MB) components, whereas hypersonic attenuation was found from the width of these components.

Molecular scattering spectra were obtained with the previously described [8] apparatus consisting of a multiple-pass Fabry–Perot interferometer and a single-mode $\lambda = 514.5$-nm argon laser. Fig. 1 shows the temperature dependence of hypersonic velocity for two different solutions. Curve A refers to the G-G solution with addition of water, in which the temperature coefficient of velocity $\mathrm{d}V/\mathrm{d}T = -6.5$ $\mathrm{m\,s^{-1}\,K^{-1}}$ above the upper critical point, and -11.6 $\mathrm{m\,s^{-1}\,K^{-1}}$ below the lower critical point. Curve B refers to the solution with addition of CCl_4, in which case $\mathrm{d}V/\mathrm{d}T = -4.8$ $\mathrm{m\,s^{-1}\,K^{-1}}$ above the upper and -11 $\mathrm{m\,s^{-1}\,K^{-1}}$ below the lower critical points. The curves A and B are virtually similar, and the temperature coefficient above the upper critical point is half that below the lower one.

Figure 2 presents the temperature dependences of hypersonic velocity for (A) a solution of pure components with no admixtures, (B) a solution with a lamination loop produced by adding water and then destroyed by adding alcohol, and (C) a solution with a double critical point. In the region of coincidence of the low and upper critical points on this curve one sees a small portion where $\mathrm{d}V/\mathrm{d}T = 0$; this portion occupies the region of the double critical point, but the rest of the curve follows the $V(T)$ dependence characteristic of the homogeneous region of solutions having closed laminated regions in their phase diagrams.

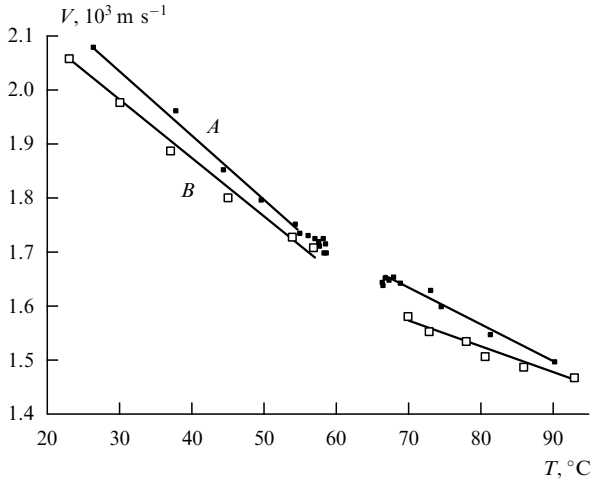

Figure 1. A, ■ — temperature dependence of the hypersonic velocity in guaiacol-glycerol solution with addition of water, with a laminated region $\Delta T = 7.28°$; B, □ — the same for guaiacol-glycerol solution with addition of CCl_4, with a laminated region $\Delta T = 1\,°C$.

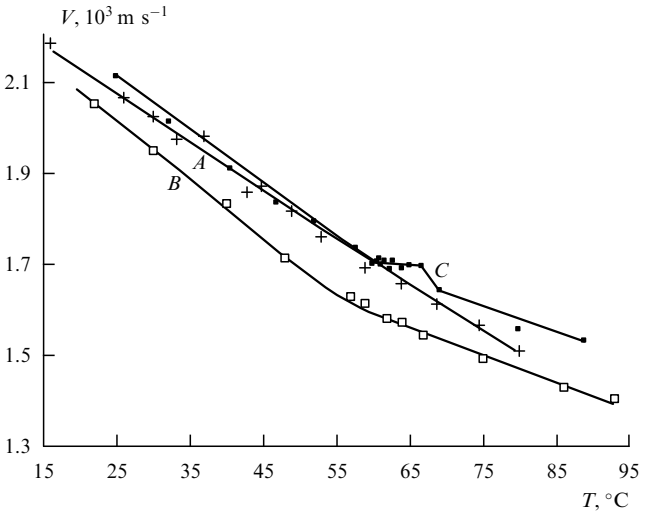

Figure 2. A, + — temperature dependence of hypersonic velocity in guaiacol-glycerol solution without additions; B, □ — the same for guaiacol-glycerol solution with addition of water, with a laminated region of 2.4 °C destroyed by adding alcohol; C, ■ — the same for guaiacol-glycerol solution with addition of water, with a double critical point.

Figures 3 and 4 present findings on hypersonic absorption in solutions with a laminated region of 7 °C and with a double critical point, respectively.

The present measurement results and qualitative observations concerning various properties of the G-G solution with admixture of negligible amounts of a third component call for a more general theoretical interpretation than they have been given before.

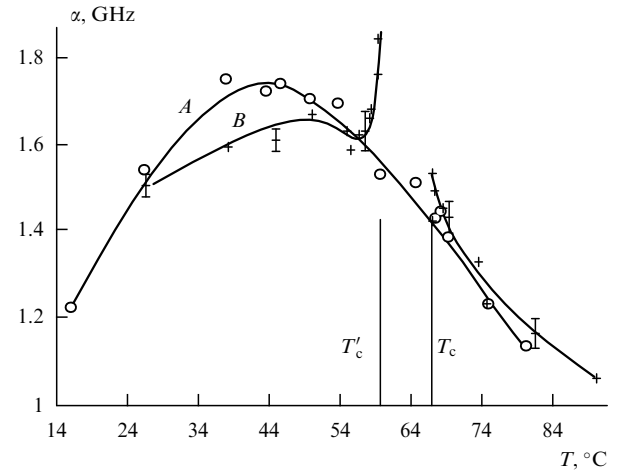

Figure 3. *A*, ○ — temperature dependence of hypersonic absorption in guaiacol-glycerol solution without additions; *B*, + — the same for guaiacol-glycerol solution with addition of water, with a laminated region of 7 °C, in which a near-critical λ-curve feature is observed.

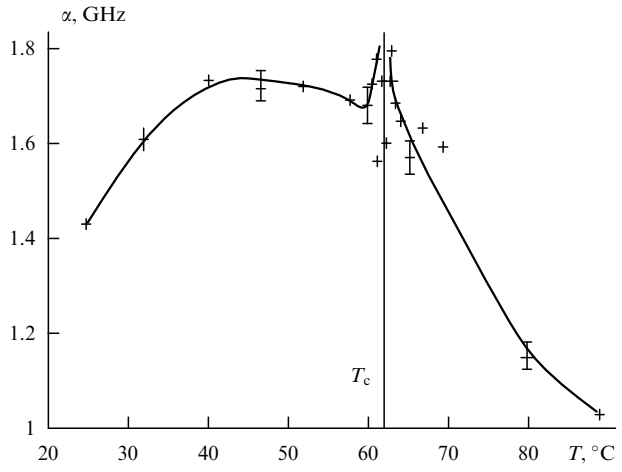

Figure 4. Temperature dependence of hypersonic absorption in guaiacol-glycerol solution with addition of water, with a double critical point; notice the λ-curve feature near this latter.

References

1. Anisimov M A *Kriticheskie Yavleniya v Zhidkostyakh i Zhidkikh Kristallakh* (Critical Phenomena in Liquids and Liquid Crystals) (Moscow: Nauka, 1987)
2. Fisher M *Priroda Kriticheskogo Sostoyaniya* (The Nature of Critical Points) (Moscow: Mir, 1968)
3. Stanley G *Fazovye Perekhody i Kriticheskie Yavleniya* (Introduction to Phase Transitions and Critical Phenomena) (Moscow: Mir, 1973)
4. Patashinskiĭ A Z, Pokrovskiĭ V L *Fluktuatsionnaya Teoriya Fazovykh Perekhodov* (Fluctuation Theory of Phase Transitions) (Moscow: Nauka, 1982)
5. Walker J S, Vause C A *J. Chem. Phys.* **79** 2660 (1983)
6. Krivokhizha S V, Fabelinskiĭ I L, Chaikov L L *Pis'ma Zh. Eksp. Teor. Fiz.* **57** 21 (1993) [*JETP Lett.* **57** 19 (1993)]
7. Krivokhizha S V et al. *Zh. Eksp. Teor. Fiz.* **103** 115 (1993) [*JETP* **76** 62 (1993)
8. Kovalenko K V, Krivokhizha S V, Fabelinskiĭ I L *Pis'ma Zh. Eksp. Teor. Fiz.* **60** 864 (1994) [*JETP Lett.* **60** 868 (1994)]
9. Samoilov O Ya *Zh. Fiz. Khim.* **52** 1857 (1978)

PACS numbers: **72.20.–i**, 75.50.Pp

Cooperative electric phenomena in degenerate magnetic semiconductors with spontaneous phase separation

E L Nagaev, V V Osipov, A A Samokhvalov

The present report is devoted to specific features of the charge transport and screening of an external electric field in degenerate magnetic semiconductors with a particular property — the phase separation in the crystal ground state. As their example, appearance of narrow high-current peaks upon application of a strong stationary electric field should be pointed out. In a certain sense this phenomenon resembles the Gunn's effect in nondegenerate nonmagnetic semiconductors though it differs from the latter in the frequency range and — what is still more important — in disappearance on exposure to a relatively weak external magnetic field. In addition, it resembles also the charge transport by the charge density waves (CDW) in quasi-one-dimensional nonmagnetic systems. But, certainly, in the case under consideration the charge transport mechanism should be quite different.

Magnetic semiconductors involve the compounds of transition or rare-earth metals displaying magnetic ordering. As an example, EuTe should be called and it will be the main subject of the report. In this compound, the spins of the Eu^{2+} ions are ordered antiferromagnetically with the Néel temperature $T_N = 9.6$ K (like in MnO, the planes perpendicular to [111] are ordered ferromagnetically while magnetic moments of adjacent planes are antiparallel) [1].

At low temperatures EuTe is virtually an insulator. But doping with the donor impurity, e.g., I or Eu in excess, causes appearance of the conductivity as in a conventional nondegenerate semiconductor. At donor densities from $\sim 10^{18}$ to 10^{19} cm^{-3}, the semiconductor becomes degenerate. This means so large overlapping of electron orbitals for neighbouring donor levels that these electrons get delocalized. The electron delocalization may be interpreted as formation of an impurity metal from impurity atoms inside the insulating crystal EuTe, in full analogy with formation of the conventional metal from the usual K or Na atoms with decreasing their interatomic distances.

In high-alloy nonmagnetic semiconductors the impurity metal appears at all temperatures, i.e. at any temperature they are high-conductive. But heavily doped magnetic semiconductors may behave quite differently at low temperatures, though at high temperatures their properties are similar to those of nonmagnetic semiconductors. Namely, at low temperatures, in spite of a large number of conduction electrons in them, they may be insulators and offer an insulator-to-metal transition as the temperature increases.

The cause of it is the magnetoelectronic phase separation in degenerate magnetic semiconductors predicted by one of the authors of the report still in 1970 [2] and described in detail in [1] and subsequent review articles, e.g., in [3]. At present this phenomenon attracts widespread attention as it influences properties of HTSC to a large extent.

Essentially, the phenomenon may be described as follows. In a magnetic semiconductor, the position of the conduction band bottom E_c depends on the type of magnetic ordering. At an energetically favoured direction of the electron spin, the conduction band bottom is lower by the quantity U at the ferromagnetic (FM) ordering than at the antiferromagnetic

(AFM) ordering. In Eu chalcogenides, this quantity amounts to 0.5 eV, that is three orders of magnitude higher than the energy D of the superexchange between magnetic atoms which is on the order of T_N.

Let us assume that conduction electrons are introduced into an AFM crystal. As their energy will be lower on the FM ordering, they tend to establish it. But direct exchange between the Eu ions hinders it (for experts in the physics of magnetism a more accurate term 'superexchange' should be used). For this reason the entire crystal can get ferromagnetic only at a very large electron density.

But at smaller electron densities the one portion of the crystal may get FM, and all the conduction electrons concentrate inside it. The other portion therewith remains AFM and devoid of conduction electrons. This is the main feature of magnetoelectronic phase separation which differs basically from the well-known phase separation at the first-order phase transition by the fact that it takes place in the ground state of the crystal. This is essentially the novel type of a crystal ground state.

To begin with, it is advisable to consider a single electron assuming that it establishes FM ordering inside a spherical region of a radius R. Then its potential energy reduces by U, and the kinetic energy enhances by a quantity proportional to $1/R^2$. In addition, replacement of the AFM ordering by the FM one increases the direct exchange energy by a quantity of order DR^3. The equilibrium value of the radius should be found from the condition of the minimum total energy E_1.

If there are other electrons in the crystals one might create a separate FM region for each of them. As an alternative, one may try to save the direct exchange energy by placing several electrons inside the same FM region. But in doing so, one arrives at an increase in the electron kinetic energy as each successive electron can be put only in a higher energy level than the preceding one. In addition, an accumulation of electrons inside the same region increases the Coulomb energy of the system. For this reason the real geometry of the phase-separated state at $T = 0$ should be found from the condition of minimum total energy consisting of the electron kinetic energy, the energy of the Coulomb interaction of the electrons with each other and with the ionized donors, and the energy spent for changing the type of the magnetic ordering in a part of the crystal. At finite temperatures, the geometry is found from the condition of minimum free energy of the system.

Results of calculations are as follows. Unlike the usual phase separation, here the two-phase state is possible only at sufficiently low temperatures. At higher temperatures the two-phase structure melts. It should also be destroyed by a magnetic field. The origin of this effect is quite obvious: the field tends to establish the FM ordering in the entire crystal, and, hence, there is no reason for the electron localization inside a certain part of the crystal. If the field is not strong enough to destroy the two-phase state it enhances the FM portion of the crystal.

The geometry of the two-phase state essentially depends on the electron number density n. For relatively small n, the FM portion of the crystal consists of separate droplets with a radius of several nm forming a superlattice inside the AFM host. This structure is perturbed, to some extent, by fluctuations of the electrostatic potential of randomly located donors. The number of electrons per droplet z and the total volume of the FM phase increase with growth in the average number density n.

At a certain critical value n_p, the FM droplets make contacts with each other, and percolation of the electron liquid and FM ordering begins. At $n > n_p$, not the FM droplets but the AFM ones form a superlattice, and the host is FM.

At $T = 0$, the configuration which realises at $n < n_p$ corresponds to an insulating state of a high-alloy semiconductor: the pinning of the droplets by the random electrostatic field of impurities makes it impossible for droplets to move over the crystal in weak electric fields. In addition, neglecting the electron tunnelling between FM droplets, one founds each electron to be locked inside its droplet, so that it cannot move over the crystal.

At finite temperatures, the electrons can be excited into the conduction band of the AFM portion of the crystal. Hence, the FM droplets play the part of donors. Their specifics consists in the fact that they are multiple-charged: in the vicinity of n_p, the number of electrons per droplet may amount to several tens. At $n > n_p$, the crystal behaves like a conventional degenerate semiconductor with a high conductivity. All these theoretical findings are confirmed by experimental investigations of doped EuTe, EuSe and other materials.

We are coming now to the results of our experimental investigation on the heavily doped EuTe in the phase-separated FM – AFM state with $n < n_p$ in strong electric fields (in part these data have been published in [4]). The results presented below were obtained for a sample with the specific resistance $\rho = 2.5 \times 10^3$ Ohm cm at 4.2 K corresponding to a nondegenerate semiconductor. On increase in T, its resistivity diminishes with the activation energy of 0.01 eV. Close to 90 K this sample acquires normal properties of a degenerate semiconductor with $\rho \sim 10^{-2}$ Ohm cm.

To avoid heating of samples, the pulse regime was used: the rectangular voltage pulses of 10 μs duration were applied to the samples. Below the threshold field of 1.4 kV/cm, the response of the system reproduces the shape of the voltage pulse following the Ohm law virtually. But after crossing a threshold, a narrow triangular peak appears at the tailing edge of the current pulse. Its height exceeds that of the ohmic background more than by a factor of 20. The current peak duration does not depend on the electric field strength and amounts to 0.7 μs.

On further increase in the field strength, the current peak shifts towards the leading edge of the pulse, and a new peak appears at the tail which also shifts forward, and then again a new peak appears at the tailing edge. The maximal number of observed current peaks runs into 5. Thus, the repetition rate for peaks increases with the field strength.

On increase in temperature, the relative height of the peaks reduces, and they disappear near 6 K, i.e. still before achievement of T_N. The peak width diminishes, too. A similar effect is observed when one applies an external magnetic field at a temperature fixed. But at 4.2 K, peaks remain observable up to maximal fields of 2 kOe used in the investigation. In both these cases the relative decrease in the peak height reflects, to a considerable extent, an increase in the ohmic component of the current.

Going over to the discussion of the mechanism for the current peak formation, one should point out that the peaks are related to the magnetic state of a sample, and more specifically, to the presence of the FM charged droplets in the sample, i.e. to the two-phase state of the sample. This follows from the fact that the same factors which destroy

these droplets (rise of temperature or magnetic field) suppress also the current peaks. Let us consider two such possible mechanisms.

The first of them is related to the N-shaped *I-V* characteristic of samples in the two-phase state. Such a shape of *I-V* characteristic is typical of nondegenerate semiconductors with multiple-charged donors, for example, Ge:Au [5, 6] or Ge:Cu [7]. In semiconductors with the N-shaped *I-V* characteristic, moving strong-field domains with a reduced conduction electron number density may arise. These domains have nothing in common with the above-mentioned phase separation: they are related to the donor ionization degree. In a certain region it is reduced, but the crystallographic and magnetic states of this region are identical to those in the rest of the crystal. This region moves over the sample from one electrode to the other. When such current domains reach the electrode and disappear at it, a current peak arises which exists up to the moment when the novel moving domain appears at the opposite electrode. The time of the domain motion between the electrodes diminishes with increase in the electric field strength, i.e. the current oscillation frequency increases with the latter [8], which resembles the regularity of current peak appearance observed by us.

Principally, such a mechanism might work in our case too, as each FM droplet is a multiple-negatively-charged donor. Capture by such a donor of an electron from the AFM conduction band should occur by surmounting the Coulomb barrier which is still more powerful than that for the multiple-charged Au donors in Ge. But it is not clear whether the current domain can move in such an imperfect crystal as a degenerate semiconductor: the impurity can pin it.

In addition, it is not clear at all whether the electrons in the AFM conduction band play an important part in the phenomenon considered, or the ohmic background is related to the electron tunnelling between the FM droplets. To distinguish between these two possibilities on the basis of their temperature dependences is very difficult as not only the ohmic current but also the tunnel current should increase with temperature. The origin for this is a reduction of the height of the potential barrier separating the droplets caused by the thermal destruction of the magnetic ordering. But in the tunnelling between the droplets, the recombination processes leading to the N-shaped *I-V* characteristic do not occur and, hence, there are no reasons for the appearance of the N-shaped *I-V* characteristic.

It should be added to this that we managed to obtain a direct proof of the fact that in several samples with a higher conductivity than discussed above ($\rho \sim 10 - 100$ Ohm cm at 4.2 K), the S-shaped *I-V* characteristics were discovered instead of the N-shaped ones. But the S-shaped *I-V* characteristic leads to the switchings instead of the current peaks. It is not self-obvious that, in samples with the closely related resistivities, the S-shaped *I-V* characteristic will be replaced by the N-shaped one. For this reason certain doubts exist whether just this mechanism realizes in our case. Perhaps, one cannot exclude that other phase-separated magnetic semiconductors have the N-shaped *I-V* characteristic, and the current oscillations may appear in them with a relatively small amplitude, without sharp peaks like in Ge:Au.

Another possibility to explain the phenomenon discovered by us is based on the fact that it resembles the charge transport by CDW in such materials as $NbSe_3$. In them, beginning from a certain threshold field which increases with a rise in the sample imperfection degree, one observes growth in the conductivity with the electric field strength. Simultaneously, noises increase sharply, and an alternating current component appears with a frequency rising with the field strength [9]. It should be mentioned that in our case the noise level remains low at field strengths above the threshold value.

Basing on the charge transport by CDW, one may assume that in our case the current peaks correspond to coherent motion of the FM droplets in the crystal which takes place as a result of their depinning by a strong electric field. In other words, the charge transport is realized by the droplet lattice moving as a whole. The coherence of the droplet motion is ensured by a strong Coulomb interaction between them. As a result, the droplets belonging to the same layer reach the cathode simultaneously, and then a current peak appears. After the layer passing the electrode, the current falls down again. The droplets from the same layer reach the cathode concurrently, and this ensures the narrowness of the peaks and a low noise level.

It is essential that all the droplets should move with the same activation energy. For this reason, if the activation energies for independent droplet motion differ from each other, the activation energy for their coherent motion may be less than those for their separate motion.

If the superlattice, on its moving as a whole, can realize the charge transport, it can also screen an external electric field. The main regularities of such screening were found within the framework of the model for the Wigner lattice, the charge of which was compensated by the uniformly distributed charge of the opposite sign. It is proved that a uniform external electric field F caused appearance of an electric field with the periodicity of the Wigner crystal inside such a system. As for the uniform field, it is screened over the length $u = F/4\pi en$, i.e. the screening length itself depends here on the field strength.

References

1. Nagaev E L *Fizika Magnitnykh Poluprovodnikov* (Physics of Magnetic Semiconductors) (Moscow: Mir, 1983)
2. Nagaev E L *Pis'ma Zh. Eksp. Teor. Fiz.* **16** 558 (1972) [*JETP Lett.* **16** 394 (1972)]; Kashin V A, Nagaev E L *Zh. Eksp. Teor. Fiz.* **66** 2105 (1974) [*Sov. Phys. JETP* **39** 1036 (1974)]
3. Nagaev E L *Usp. Fiz. Nauk* **165** 529 (1995) [*Phys. Usp.* **38** 497 (1995)]
4. Osipov V V, Samokhvalov A A, Nagaev E L *Z. fur Physik B* **96** 31 (1994); *Pis'ma Zh. Eksp. Teor. Fiz.* **59** 788 (1994) [*JETP Lett.* **59** 819 (1994)]
5. Ridley B K, Pratt R G *Phys. Lett.* **4** 300 (1963)
6. Kurova I A, Kalashnikov S G *Fiz. Tverd. Tela* **5** 3224 (1963)
7. Zhdanova N G, Kagan M S, Kalashnikov S G *Fiz. Tverd. Tela* **8** 788 (1966)
8. Bonch-Bruevich V L, Zvyagin I P, Mironov A G *Domennaya Elektricheskaya Neustoichivost' v Poluprovodnikakh* (Domain Electrical Instability in Semiconductors) (Moscow: Nauka, 1972) [Translated into English (New York: Consultant Bureau, 1975)]
9. Fleming R, Grimes C *Phys. Rev. Lett.* **42** 1423 (1979)

PACS numbers: 03.65.Bz

The quantum theory of wave packet scattering, the causality principle, and superlight tunnelling

L A Khalfin

Although the development of the principles of quantum theory dates back 70 years or more, only in recent years have its most striking — and, from the classical viewpoint,

even paradoxical — predictions become the subject of intensive experimental study. Suffice it to mention the famous Einstein–Podolsky–Rosen (ERP) effect [1], which have received attention of new generations of physicists following the well-known work of Bell [2] (see also [3]).† More generally, the focus is on *macroscopic quantum effects*, i.e. the predictions of quantum theory in the region where classical theory was thought to have complete dominion (see, e.g., [3] and references therein). The subject of this work is one further surprising prediction of quantum theory, specifically associated with the fundamental principle of causality.

1. Einstein's causality principle, which excludes signal transmission with a velocity exceeding the speed of light, provided a basis for the theory of relativity and had been formulated within the hypotheses of classical physics before quantum theory was developed. Formally, its application in quantum theory depends on the requirement of commutation for physical operators at points separated by a space-like interval. The question arises, however, as to the informal, substantive validity of this principle within the framework of quantum theory. This question has already emerged in the discussion of the EPR effect, which was suspected (incorrectly, as was proved) of violating the causality principle on a macroscopic scale. Clearly, the analysis of the causality principle within a quantum theory requires a study of nonstationary quantum-theoretical problems, but just they are known to differ strikingly from nonstationary classical problems. Suffice it to mention that in quantum theoretical nonstationary problems it is no longer energy but rather its distribution (!) which is conserved (fixed-energy states, alias stationary states rule out any dynamics in a quantum theory). Notice also the quantum Zeno effect, which is paradoxical from the classical point of view (see [4] and references therein) and has been dubbed 'Man's made Physics' by the present author [5].

The question of the substantive validity of the causality principle in quantum theory came under study in 1932 (before the discovery of the EPR effects) in connection with two specifically quantum-theoretical phenomena, namely, the decay and formation of excited states of spatially separated atoms (Fermi [6]), and nonstationary tunnelling (MacColl [7]); and also in 1955, in the Wigner's quantum scattering theory [8]. Although the results obtained in these studies were consistent with the classical causality principle, further analysis showed them to be conditioned on certain approximations. Clearly, a principle of such fundamental importance must be investigated with exact methods free of any approximations. In fact, exact analysis showed that both in the Fermi problem (see [9] and references therein) and in the tunnelling — or scattering — problem (tunnelling is a special case of the one-dimensional scattering) [10], the classically formulated causality principle is evidently violated. It is to be emphasised at once that, in quantum theory, it is not the limitation of signal velocity but rather the *very formulation of the causality problem* which turns out to be invalid (see below). As far as the velocity limitation is concerned, not inconsistent with this particular consequence of the causality principle is, along with the EPR effect, the experimentally studied [11] superlight tunnelling, which is a special case of the classical causality 'violation' predicted in [10]. At the heart of the exact treatments referred to is the inclusion of the spectral principle of quantum theory (the existence of the vacuum state), or more specifically of the fact that the spectrum of the Hamiltonian (total energy) operator is semibounded from below

$$\operatorname{Spec} H \geqslant E_0, \quad |E_0| < \infty . \tag{1}$$

This, in turn, immediately implies that quantum theory is fundamentally *nonlocal*, which is precisely what the experimental detection of the EPR effect and superlight tunnelling have demonstrated. These, however, do not exhaust, and indeed there are even more amazing, quantum-theoretical predictions apparently paradoxical from the classical point of view.

2. Our analysis is based on the following mathematical facts. By and large let $f(E)$ be a complex-valued function from $L_1(-\infty, \infty)$, so that its Fourier transform

$$F(t) = \int_{-\infty}^{\infty} f(E) \exp(-\mathrm{i}Et)\, \mathrm{d}E . \tag{2}$$

is defined.

Proposition 1. If $f(E)$ has a discontinuity of the derivative $\mathrm{d}^n f(E)/\mathrm{d}E^n$ at $E = E_1$, then the asymptotic behaviour of $F(t)$ as $|t| \to \infty$ is determined by this discontinuity, so that

$$\left|F(t)\right|_{|t|\to\infty} \sim |t|^{-(n+1)} . \tag{3}$$

If $f(E)$ is semifinite

$$f(E) = 0, \quad E < 0 , \tag{4}$$

then it holds true

Proposition 2 (Titchmarsh's theorem [12]): $F(t)$ has an analytic continuation to the lower half $\operatorname{Im} t < 0$ complex plane, i.e. $F(t)$ is the boundary value of an analytical function in the $\operatorname{Im} t < 0$ half-plane.

If a semifinite $f(E)$ is integrable in module in the quadrant $f(E) \in L_2[0, \infty)$, then it holds true the following fundamental

Proposition 3 (Paley-Wiener theorem [13]): $F(t)$ must necessarily satisfy the condition

$$\left| \int_{-\infty}^{\infty} (1 + t^2)^{-1} \ln \left|F(t)\right| \mathrm{d}t \right| < \infty . \tag{5}$$

If $f(E) \in L_1(-\infty, \infty)$ is nonnegative: $f(E) \geqslant 0$, $E \in (-\infty, \infty)$, then it holds true

Proposition 4 [14]: If $f(E)$ offers a finite first moment

$$\left| \int_{-\infty}^{\infty} E f(E)\, \mathrm{d}e \right| < \infty , \tag{6}$$

then

$$\left. \frac{\mathrm{d}\left|F(t)\right|^2}{\mathrm{d}t} \right|_{t=0} = 0 . \tag{7}$$

If $f(E) \in L_1(-\infty, \infty)$ is nonnegative and normalized to unity

$$\int_{-\infty}^{\infty} f(E)\, \mathrm{d}E = 1 , \tag{8}$$

† Recent harnessing of ideas (pertinent to EPR effect) for launching the quantum cryptography showed once again the surprises at applications of fundamental research findings which appeared as academic only.

then $F(t)$ is a characteristic function (in the sense of perturbation theory), which is positive definite from the Bochner–Khinchin theorem and possesses a whole series of specific nontrivial properties [15] different from the general properties of the Fourier transform.

Notice that the above properties of $F(t)$, particularly (5), provide an important refinement for the familiar uncertainty relation.

The properties of $F(t)$ given above have been used in the author's previous work on various subjects, including the quantum theory of unstable physical systems (particles) (see [5, 14, 16] and references therein); the quantum Zeno effect [4, 5]; the problem of CP invariance (see [14, 16, 17] and references therein); the quantum theory of wave packet scattering and causality [10]; the justification of statistical physics [3, 18], and, finally, the analysis of the high-energy asymptotic behaviour of scattering amplitudes and vertex functions [19, 20].

The theorem from [10] has previously been formulated, proved, and used in the author's works [3 – 5, 14, 16 – 20].

3. In discussing the space and time theory of wave packet scattering, we will follow the analysis by Wigner [8]. We will avoid technical details and consider, as in [10], the s-scattering of a relativistic particle, so that $E = k$. For large enough $r > R$, where R is the interaction region, the wave function may be written in the form

$$\left.\begin{aligned} \psi(r,t) &\approx r^{-1}\left[\phi_{\text{in}}(r,t) + \phi_{\text{out}}(r,t)\right], \\ \phi_{\text{in}}(r,t) &= \int_0^\infty C(k)\exp\left[-\text{i}k(r+t)\right]\mathrm{d}k, \\ \phi_{\text{out}}(r,t) &= \int_0^\infty C(k)\exp\left[2\text{i}\delta(k) + \text{i}k(r-t)\right]\mathrm{d}k, \end{aligned}\right\} \quad (9)$$

where $C(k)$ is the wave packet momentum (energy) distribution, and $\delta(k)$ is the scattering phase. The classical notion of causality implies that $\phi_{\text{out}}(r,t)$ at $r > 1$ is negligibly (exponentially) small for sufficiently localized packets [sufficiently 'broad' $C(k)$], i.e. there is no leading effect. In accordance with this principle, it was shown in [8] that for scattering in the resonance region

$$\exp\left[2\text{i}\delta(k)\right] = \frac{k - m - \text{i}\Gamma}{k - m + \text{i}\Gamma} \quad (10)$$

$\phi_{\text{out}}(r,t)$ at $r > t$ is exponentially small (the argument of the exponential depends on the 'width' of $C(k)$), and at $r < t$, $\phi_{\text{out}}(r,t)$ produces retardation with a resonance lifetime Γ^{-1}. However, as was shown in [10], this result is inaccurate because in the calculation of the integrals in Eqn (9) in [8], only the pole singularities were included. The exact calculation of these integrals requires a detailed knowledge of $C(k)$ and $\delta(k)$. However, based on Propositions 2 and 3 (see Section 2) it can be proved that $\phi_{\text{out}}(r,t)$ in the superlight region $r > t$ cannot be equal to zero, which means the *violation of the causality principle in its classical formulation*. Furthermore, on macroscopical scales $r \gg t$, $r > R$ the asymptotic behaviour of $\phi_{\text{out}}(r,t)$ is, based on Proposition 1, determined by the discontinuity in $C(k)$ or its derivatives, and also by $\exp\left[2\text{i}\delta(k)\right]$. Thus, if the nth derivative has a discontinuity at $k = 0$, and if $\exp\left[2\text{i}\delta(k)\right]$ in this vicinity behaves like $k^{-1/2}$, then

$$\phi_{\text{out}}(r,t)_{r \gg t} \sim b_n (r-t)^{-(n+5/2)}, \quad (11)$$

where b_n is defined via the behaviour of $C(k)$ and $\delta(k)$ in the vicinity of $k = 0$. Although (11) looks like a gross violation of the classical causality principle (leading effect), this does not in reality violate Einstein's causality principle, namely, the exclusion of superluminal motion. This is due to the fact that for $r \gg t$ there is also a *nonzero contribution* from $\phi_{\text{in}}(r,t)$ (!):

$$\phi_{\text{in}}(r,t)_{r \gg t} \sim a_n (r+t)^{-(n+1)}. \quad (12)$$

Since $C(k)$ and n are at our disposal, Eqs (11) and (12) allow us to make the 'violation' of (macro)causality arbitrarily small. Physically, this 'violation' can be explained by the fact that the spectral principle (1) together with Propositions 1 to 3 make it impossible to form $\phi_{\text{in}}(r,t)$ as a *local* packet. Interestingly, results like Proposition 4 suggest that the nature of microcausality violations, i.e. the behaviour of $\phi_{\text{out}}(r,t)$ as $(r-t) \to 0+$, is determined by the behaviour of the scattering phase at infinite energies.

Now notice that by far a more significant 'violation' of causality results from including the field quantum-theoretical effects. In fact, by the optical theorem the forward-scattering amplitude (or, more accurately, its imaginary part) at a given energy (momentum) $E = k$ is determined by allowable inelastic processes, namely, by multiple particle production events. As a result, $\exp\left[2\text{i}\delta(k)\right]$ is discontinuous at the corresponding $E = E_l$ thresholds, whose number in the region $E \in (0, \infty)$ is infinite. But then, if the behaviour of $\exp\left[2\text{i}\delta(k)\right]$ near the $E = E_l$ threshold is $(E - E_l)^{1/2}$, this gives the following contribution to $\phi_{\text{out}}(r,t)$ at $r \gg t$:

$$\phi_{\text{out}}(r,t) \sim d_l (r-t)^{-3/2}, \quad (13)$$

whereas $\phi_{\text{in}}(r,t)$ is determined by the behaviour of $C(k)$ at $k = 0$ [cf. (12)] and does not depend on the behaviour of the inelastic processes. The violation of classical causality (macrocausality) discovered in [10] is investigated in a similar manner for more general elastic scattering processes and for inelastic processes.

Experimentally, the observation of the (macro)causality 'violation' effects predicted for ordinary elementary particles is complicated by the fact that no coherent sources of wave packets (i.e. no elementary-particle 'lasers') are available (the separate paper will be devoted to studies on coherent wave packet formation in the accelerator experiments). Fortunately, they are available for photons, and it is this circumstance which leads to the discovery of superlight tunnelling [11].

4. One-dimensional tunnelling is a special case of the scattering we have discussed above. Moreover, for photons this is a relativistic region in which $E = k$. The familiar methods for treating the tunnelling problem, which Gamow [21] used to explain the α-decay, involve the quasiclassical picture of stationary tunnelling [22]. The quasiclassical approach yields the expression for stationary tunnelling under the assumption that the potential barrier is sufficiently thin.

The nonstationary tunnelling of a wave packet through a potential barrier of a general shape cannot be treated quasiclassically because the accuracy in the value of the stationary tunnelling rate in this approximation depends significantly and nonuniformly on the energy. As for potential barriers amenable to exact stationary tunnelling solutions (say, a rectangular barrier [23]), they involve artifacts which exclude their use in the nonstationary case.

Notice that the disagreement (see [23] and references therein) as to the time a tunnelling particle spends inside the barrier (an important question for the Josephson effect and the tunnelling microscope) has its explanation both in that quantum theory includes no time operator and that the quasiclassical stationary approximation has been used. In actual fact, as the argument above has shown, *the very question about the time the particle is inside the barrier is incorrect*, whereas the rigorous methods of Sections 2 and 3 provide an exact answer to a specific problem of nonstationary wave packet tunnelling.

The exact results of Sections 2 and 3 imply the 'violation' of causality principle in nonstationary (one-dimensional) tunnelling of photons (superlight tunnelling). This unexpected and, from the classical viewpoint, paradoxical phenomenon (note the paper title 'When Can Light Go Faster Than Light?'[24]) was recently observed in brilliant experiments at Berkeley (see [11] and references therein). This type of experiment has become possible due to recent advances in quantum optics, namely the realization of potential barriers for photons (quarter-wave layered dielectric mirrors, or 'photonic band gaps' [25]), and the use of Hong-Ou-Mandel (HOM) two-photon interferometers [26] in subfemtosecond twin-photon coincidence schemes based on the two-photon decay of uv laser photons in $\chi^{(2)}$-nonlinear crystals.

It is essential from the viewpoint of quantum theory that the *summary* energy of photon twins is fixed by the 'parent' energy of uv laser photon, whereas the scheme of subfemtosecond coincidence permits the determination of time *difference* for photon twins arriving in the coincidence scheme of a HOM interferometer. Fixing of energy sum and the time difference for photon twins is not in contradiction with the uncertainty relation, as also there exists no contradiction between the latter and the result of fixing the coordinate sum and the momentum difference in the EPR effect [1].

The experimental arrangement used and the results obtained were briefly outlined in the verbal communication.

References

1. Einstein A, Podolsky B, Rosen N *Phys. Rev.* **47** 777 (1935)
2. Bell J S *Physics* **1** 195 (1964)
3. Khalfin L A, Tsirelson B S *Found. Phys.* **22** 879 (1992)
4. Khalfin L A *Usp. Fiz. Nauk* **160** 185 (1990) [*Sov. Phys. Usp.* **33** 868 (1990)]
5. Khalfin L A — invited lectures at Bell Laboratory and National Institute of Standards and Technology at Boulder, CO, November 1990
6. Fermi E *Rev. Mod. Phys.* **4** 87 (1932)
7. MacColl L A *Phys. Rev.* **40** 621 (1932)
8. Wigner E P *Phys. Rev.* **98** 145 (1955); Goldberger M, Watson K *Collision Theory* (New York: Wiley, 1964) [Translated into Russian (Moscow: Mir, 1967)]
9. Hegerfeldt G C *Phys. Rev. Lett.* **72** 596 (1994)
10. Khalfin L A *Pis'ma Zh. Eksp. Teor. Fiz.* **11** 46 (1970) [*JETP Lett.* **11** 28 (1970)]
11. Steinberg A M, Chaio R Y *Phys. Rev. A* **51** 3525 (1995); Chaio R Y, Kwait P G, Steinberg A M *Quantum Nonlocality in Two-Photon Experiments at Berkeley*. Preprint www http://xxx.lanl.gov/archive/quantph/9501016
12. Titchmarsh E C *Introduction to the Theory of Fourier Integrals* (Oxford: Clarendon Press, 1937) [Translated into Russian (Moscow: GITL, 1948)]
13. Paley R E A C, Wiener N *Fourier Transforms in the Complex Domain, Amer. Soc. Coll.* **XIX** (1934)
14. Khalfin L A *Pis'ma Zh. Eksp. Teor. Fiz.* **8** 106 (1968) [*JETP Lett.* **8** 65 (1968)]
15. Doob J *Stochastic Processes* (New York: Wiley, 1953) [Translated into Russian (Moscow: GIIL, 1956)]
16. Khalfin L A *Dokl. Akad. Nauk SSSR* **115** 277 (1957); *Zh. Eksp. Teor. Fiz.* **33** 1371 (1958) [*Sov. Phys. JETP* **6** 1053 (1958)]; *Kvantovaya Teoriya Raspada Fizicheskikh Sistem* (Quantum Theory of Decay of Physical Systems) Candidate's Dissertation Thesis (Moscow: FIAN, 1960); *Issledovaniya po Kvantovoi Teorii Nestabil'nykh Elementarnykh Chastits* (Studies on the Quantum Theory of Unstable Elementary Particles) Doctor's Dissertation Thesis (Moscow: FIAN, 1968: Dubna: LTF JINR, 1972)
17. Khalfin L A Report at the scientific session of the Division of Nuclear Physics of the Academy of Sciences of the USSR (1968); Trudy Seminara "Teoretiko-gruppovye Metody v Fizike" (Proceeding of the Seminar 'Group-Theoretical Methods in Physics') (Ed. M Markov) Vol. 2 (1986) p. 608; "Teoriya $K^0-\bar{K}^0$, $D^0-\bar{D}^0$ and $B^0-\bar{B}^0$ Mezonov vne Priblizheniya Vaiskopfa-Vignera i Problema *CP*-Invariantnosti" (Theory of $K^0-\bar{K}^0$, $D^0-\bar{D}^0$ and $B^0-\bar{B}^0$ Mesons beyond the Weisskopf-Wigner Approximation, and the Problem of *CP* Invariance) Preprint LOMI P-4-80 (Leningrad, 1980); Khalfin L A "New results on the *CP*-violation Problem", Preprint CPT, DOE-ER 40 200-211 (Austin: University of Texas at Austin, 1990); "A New CP-violation Effect and a New Possibility for Investigations of K_s, $K_L(K^0, K^0)$ Decay Modes", Preprint CPT, DOE-ER 40 200-246 (Austin: University of Texas at Austin, 1991); "*CP*-violation and *CPT*-invariance Tests Beyond Standard Lee–Ohme–Yang Theory", II Eurodafne Collaboration Meeting, Frascati, 19–22 April, 1994, p. 249, INFN, LNF-94/033(IR), (Frascati, 1994); Khalfin L A "*CP*-violation Problem Beyond Standard Lee–Ohme–Yang Theory", June 1994, submitted to "The DAFNE Physics Handbook"; "On the *CPT*-invariant Theory of *CP*-violation", Preprint PDMI-3/1995, February 1995
18. Khalfin L A *Dokl. Akad. Nauk SSSR* **162** 1273 (1965)
19. Khalfin L A *Zh. Eksp. Teor. Fiz.* **45** 631 (1963) [*Sov. Phys. JETP* **18** 433 (1964)]
20. Khalfin L A *Pis'ma Zh. Eksp. Teor. Fiz.* **15** 215 (1972) [*JETP Lett.* **15** 149 (1972)]
21. Gamow G *Z. Physik* **51** 204 (1928)
22. Landau L D, Lifshitz E M *Quantum Mechanics. Nonrelativistic Theory* (New York: Pergamon Press, 1967)
23. Steinberg A M *Phys. Rev. Lett.* **74** 2405 (1995)
24. Steinberg A M *When Can Light Go Faster Than Light?, The Tunnelling Time and its Sub-femtosecond Measurement via Quantum Interference*, Ph. D. Thesis (Berkeley: California University, 1994)
25. Yablonovitch E, Leung K M *Physica B* **175** 81 (1991)
26. Hong C K, Ou Z Y, Mandel L *Phys. Rev. Lett.* **59** 2044 (1987)