Doklady Physics, Vol. 49, No. 12, 2004, pp. 697–700. Translated from Doklady Akademii Nauk, Vol. 399, No. 4, 2004, pp. 468–471.

PHYSICS

Dynamics of Virtual-Cathode Formation in a Viscous-Friction Medium

A. E. Dubinov

Presented by Academician V.V. Osiko June 25, 2004

Received June 25, 2004

It is well known that the spatial charge of an electron beam that ensures a sufficiently high current and that is injected into an equipotential cavity forms a potential barrier within the cavity that prevents further electron injection. In this case, some of the electrons are reflected by the potential barrier backward to the point of their injection, and the turning point, which can be characterized by the infinite electron concentration in the hydrodynamic limit, is called a *virtual cathode* (VC). In recent years, electron devices with a VC have been commonly used in vacuum high-current electronics. On this basis, high-power microwave and x-ray sources, collective ion accelerators, high-current switches, and high-vacuum meters were proposed. The current status of the physics and technology of VC-based devices was reviewed in [1, 2].

At present, the general dynamics of formation and existence of VCs in electron beams cannot be rigorously analyzed, because the available methods are inapplicable in the case of a multivalued velocity. Therefore, computer simulation is the basic method for the theoretical analysis of the above systems. Unsteady VC-formation dynamics was apparently first considered in [3]. Nevertheless, the initial stage of VC formation in vacuum can be self-consistently analyzed for simple cases, such as a half-space [4–6], a plane equipotential cavity [6], and a cylindrical cavity with radial injection [7]. It is of practical interest to consider the features of VC-formation dynamics for the equipotential cavity filled with a certain medium. For example, VC-formation dynamics in a dielectric medium in the absence of both permittivity dispersion and electron scattering was analytically analyzed in [8].

In this paper, we present the analytical self-consistent solution to the problem of the initial stage of VC formation in the case of the injection of the beam into a half-space filled with a model medium, where electrons are scattered. The averaged action of this scattering is described by the friction force that is proportional to the electron velocity. This friction acting on electrons in the medium is similar to the viscous friction that acts during the motion of bodies in a liquid and that can arise under certain conditions, for example, when free electrons move in semiconductors [9, 10] and gases [11].

Thus, we consider the injection of a monoenergetic electron beam into the half-space that is bounded by the metallic plane $z = 0$ and that is filled with a medium in which electrons are subjected to friction proportional to their velocity. In this case, it is not important whether the electrons are injected through the thin metallic electrode that is transparent to them (in this case, a positive image charge is induced on the electrode) or whether the electrons penetrate into the half-space by means of being knocked out from the surface, e.g., by an intense photon flux (in this case, an uncompensated charge of positive ions is left on the photoemitter surface). In both cases, the positive charge is exactly equal to the total charge of the electron beam at any time.

Let electrons be emitted from the surface $z = 0$ with a beam density $N(t)$ cm^{-2} s^{-1} at $t = 0$. Until the time t, the electron space charge density is equal to

$$Q = -e\int_0^t N(\tau)d\tau, \qquad (1)$$

and the positive surface charge is equal to $|Q|$. Therefore, in accordance with [3, 5], within the injection-time interval from τ to $\tau + d\tau$ the electrons are subjected to an electric field $E(\tau)$ whose magnitude is determined by the charge of electrons emitted previously. Therefore,

$$E(\tau) = 4\pi e\int_0^\tau N(\tau)d\tau. \qquad (2)$$

Relation (2) is valid only up to the time when a certain electron layer outruns a layer injected previously. Below, we derive the exact expression for this time. For

Sarov State Physicotechnical Institute, Sarov, Nizhegorodskaya oblast, Russia
e-mail: dubinov@ntc.vniief.ru; dubinov@rol.ru

1028-3358/04/4912-0697$26.00

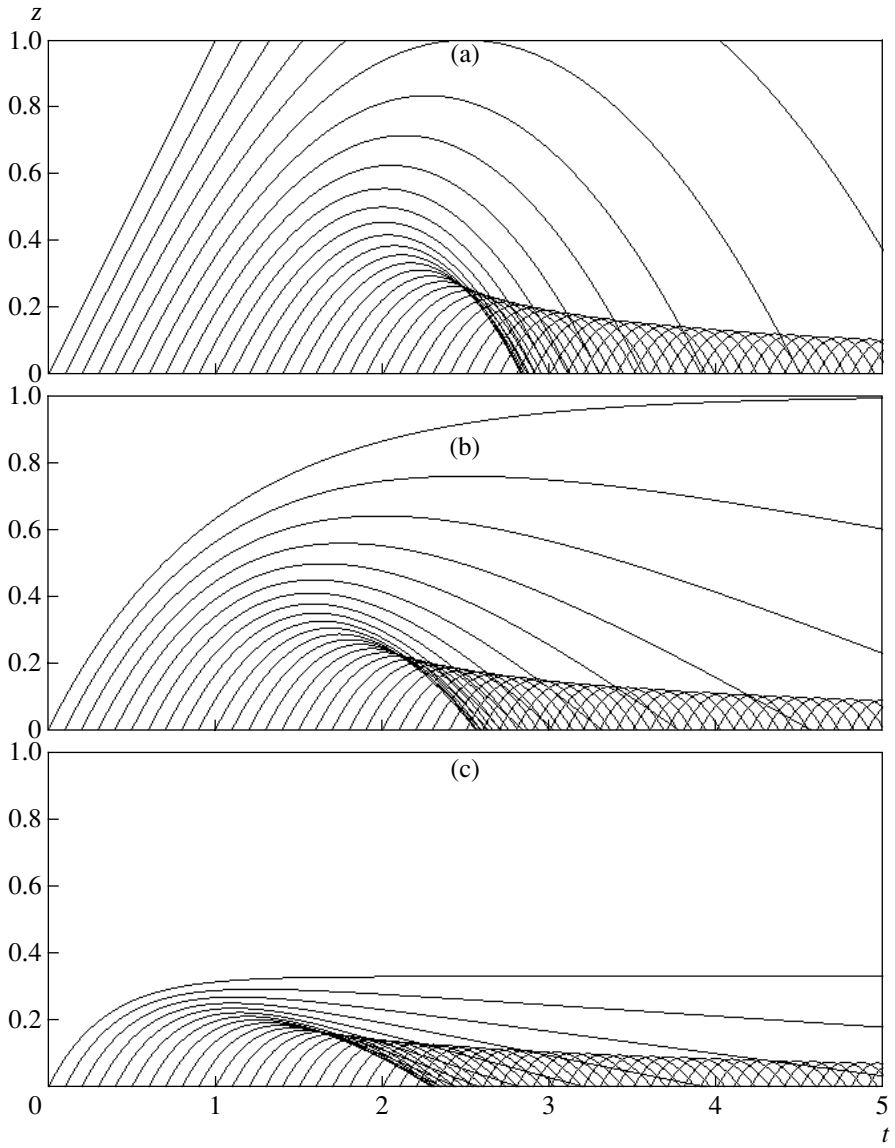


Fig. 1. Electron trajectories in the (z, t) plane for k = (a) 0, (b) 1, and (c) 3.

a rectangular pulse with a density $N(t) = N_0\Theta(t)$, where $\Theta(t)$ is the Heaviside step function, we find

$$E(\tau) = 4\pi e N_0 \tau. \quad (3)$$

Then, the equation of motion for electrons in terms of Lagrangian variables (t, τ) is written in the form

$$\ddot{z}(t, \tau) = -\frac{e}{m}E(\tau) - k\dot{z}(t, \tau) = -\frac{4\pi e^2 N_0}{m}\tau - k\dot{z}(t, \tau) \quad (4)$$

with the initial conditions

$$z(\tau, \tau) = 0, \quad \dot{z}(\tau, \tau) = v_0. \quad (5)$$

Moving on to the dimensionless variables in terms of the time unit $\tau' = \dfrac{mv_0}{4\pi e^2 N_0}$, the length unit $v_0\tau'$, the velocity unit v_0, the electron-density unit N_0, the electric-field unit $\dfrac{mv_0}{e\tau'}$, the potential unit $\dfrac{mv_0^2}{e}$, etc. (below, primes are omitted), we transform Eqs. (4) and (5) into the form

$$\ddot{z}(t, \tau) = -\tau - k\dot{z}(t, \tau), \quad z(\tau, \tau) = 0, \quad \dot{z}(\tau, \tau) = 1. \quad (6)$$

The solution of problem (6) is easily obtained in the form

$$v(t, \tau) = \frac{(\tau + k)\exp[-k(t - \tau)] - \tau}{k} \quad (7)$$
$$\{v(t, \tau) = 1 + \tau^2 - t\tau\};$$

$$z(t,\tau) = \frac{(\tau+k)\exp[-k(t-\tau)] - (\tau+k) + \tau k(t-\tau)}{k^2}$$

$$\left\{ z(t,\tau) = t - \tau + t\tau^2 - \frac{t^2\tau}{2} - \frac{\tau^3}{2} \right\}. \quad (8)$$

Below, the braces on the right-hand sides of formulas enclose the corresponding expressions for vacuum conditions from [6]. Moreover, it is also easy to verify that the expressions in braces are obtained in the limit $\lim_{k\to 0}$.

Figure 1 shows calculated electron trajectories in the (z, t) plane for various parameters k. These trajectories clarify the role of friction: the larger the parameter k, the earlier and closer to the injection plane the VC is formed.

The turning time t_r for different electron layers ($v = 0$) is determined from the condition

$$t_r = \frac{\tau k - \ln\dfrac{\tau}{\tau+k}}{k} \quad \left\{ t_r = \tau + \frac{1}{\tau} \right\}, \quad (9)$$

where the time t_p of the return of electrons to the plane can be determined from the equation $z(t_p, \tau) = 0$ (see Eq. (8)). This equation has two roots:

$$t_p = \tau,\ t_p = \frac{\tau W_0\left[-\dfrac{\tau+k}{\tau}\exp\left(-\dfrac{\tau+k}{\tau}\right)\right] + \tau^2 k + \tau + k}{\tau k}$$

$$\left\{ t_p = \tau, \quad t_p = \tau + \frac{2}{\tau} \right\}, \quad (10)$$

where $W_0(x)$ is the basic real branch of the Lambert function [12] and the first and second roots correspond to electron injection and to the return of electrons to the $z = 0$ plane, respectively.

We now find the time when the velocity uniqueness in the electron flow is violated, i.e., the time when equality (2) becomes invalid. In the uniqueness region, we have

$$\frac{\partial}{\partial\tau} z(t,\tau) < 0. \quad (11)$$

This condition is fulfilled when

$$t_m < \frac{W_{-1}[(-1-\tau k-k^2)\exp(-1-\tau k)] + 2\tau k + 1}{k},$$

$$t_m > \frac{W_0[(-1-\tau k-k^2)\exp(-1-\tau k)] + 2\tau k + 1}{k}$$

$$\left\{ t_m < \tau\left(2 - \sqrt{1-\frac{2}{\tau^2}}\right), \quad t_m > \tau\left(2 + \sqrt{1-\frac{2}{\tau^2}}\right) \right\}. \quad (12)$$

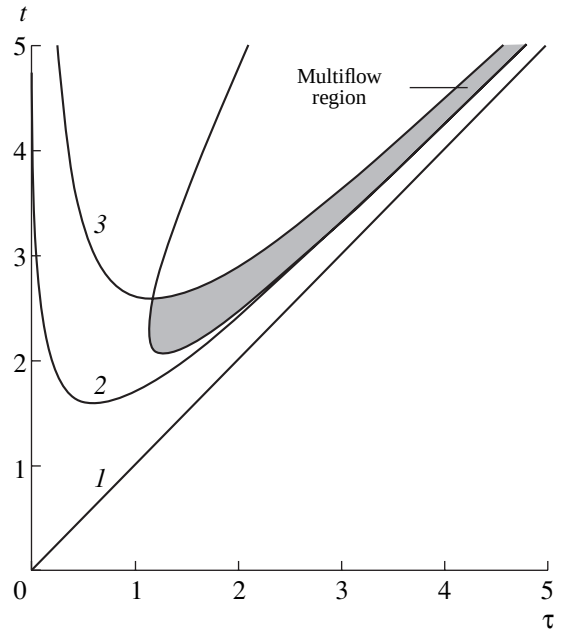


Fig. 2. Regions of the electron-beam states in the (t, τ) plane for $k = 1$.

In the first inequality of set (12), $W_{-1}(x)$ denotes the other real branch of the Lambert function, which takes values smaller than –1.

Relationships (9), (10), and (12) determine the electron-beam states whose regions in the (t, τ) plane are shown in Fig. 2. The region of the forward electron motion is bounded by straight line *1* (injection line) and curve *2* (line of the stopping and reflection of electrons), whereas the region of the backward electron motion is bounded by curves *2* and *3* (the latter is the line of electron incidence on the plane). The dark region is the multivalent-velocity (multiflow) region, where basic expression (2) for the electric field is violated and our model is invalid. However, the boundaries of the multiflow region and the dynamics in the uniqueness region remain valid even after the formation of the VC and multiflow regime.

We now determine the duration t_{VC} of the VC-formation process as the minimum of function $t_r(\tau)$ (9). Minimization yields

$$t_{VC} = -\frac{k-\sqrt{k^2+4}}{2} - \frac{1}{k}\ln\frac{k-\sqrt{k^2+4}}{k+\sqrt{k^2+4}} \quad \{t_{VC} = 2\}. \quad (13)$$

The VC-formation time as a function of the parameter k is shown in Fig. 3.

In summary, the self-consistent analytical model for the initial stage of the VC-formation process in a medium, where electrons move with friction proportional to their velocity, has been developed and analyzed. We should also note that the Lambert function

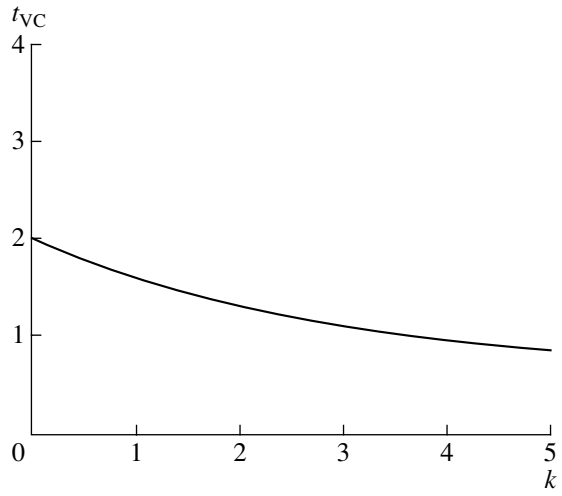


Fig. 3. Virtual-cathode formation time t_{VC} vs. the parameter k.

used above is employed in many fields of mathematics and physics. However, its properties are not described in the Russian scientific literature devoted to special functions. A comprehensive review of the properties and usage of the Lambert function can be found in [12].

ACKNOWLEDGMENTS

I am deeply grateful to Prof. G.H. Connet (Zurich, Switzerland), who kindly placed paper [12] at my disposal, and to Prof. A.A. Rukhadze (Moscow, Russia) for consultations on the problems of VC dynamics.

REFERENCES

1. A. E. Dubinov and V. D. Selemir, Radiotekh. Élektron. (Moscow) **47** (6), 645 (2002).
2. A. E. Dubinov, I. Yu. Kornilova, and V. D. Selemir, Usp. Fiz. Nauk **172** (11), 1225 (2002).
3. A. M. Ignatov and A. A. Rukhadze, Kratk. Soobshch. Fiz., No. 11, 13 (1977).
4. J. W. Poukey and N. Rostoker, IEEE Trans. Plasma Sci. **13** (10), 897 (1971).
5. N. D. Naumov, Dokl. Akad. Nauk **359** (3), 323 (1998) [Dokl. Phys. **43**, 154 (1998)].
6. A. E. Dubinov, Fiz. Plazmy **26** (5), 439 (2000) [Plasma Phys. Rep. **26**, 409 (2000)].
7. N. D. Naumov, Fiz. Plazmy **25** (2), 181 (1999) [Plasma Phys. Rep. **25**, 160 (1999)].
8. A. E. Dubinov and I. Yu. Kornilova, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 6, 71 (2001).
9. C. K. Birdsall and W. B. Bridges, *Electron Dynamics of Diode Regions* (Academic, New York, 1966).
10. P. V. Akimov and H. Schamel, J. Appl. Phys. **92** (3), 1690 (2002).
11. R. W. Schmieder, J. Appl. Phys. **50** (2), 712 (1979).
12. R. M. Corless, G. H. Gonnet, D. E. J. Hare, *et al.*, Adv. Comput. Mat. **5**, 329 (1996).

Translated by G. Merzon

Doklady Physics, Vol. 49, No. 12, 2004, pp. 701–705. Translated from Doklady Akademii Nauk, Vol. 399, No. 4, 2004, pp. 472–476.

PHYSICS

Dynamics of Formation of the Pulsed Electrical Breakdown of Highly Overvoltaged Gas Gaps in the Subnanosecond Range

S. N. Ivanov

Presented by Academician G.A. Mesyats March 24, 2004

Received May 24, 2004

The mechanisms of the initiation of pulsed electrical breakdown of gas gaps in the subnanosecond range are of great interest for gas-discharge physics. Recently developed voltage-pulse generators make it possible to form high-voltage pulses with 200-ps-wide fronts and to apply voltages much higher than the static-breakdown voltage to a gas gap. Under these conditions, field emission from a cathode with the subsequent explosion of microinhomogeneities on its surface begins to make an increasing contribution to the initiation of breakdown [1]. Moreover, for high ratios $\frac{E}{p}$ of the electric field to pressure at the stage of breakdown delay and the initial stage of commutation, the energy gained by some free electrons per unit path can be higher than the energy lost in inelastic collisions. This regime is called the regime of the continuous acceleration of electrons, and these electrons are named fast or runaway electrons [1]. Fast electrons intensely ionize the gas medium in the interelectrode gap, and the breakdown formation time can decrease sharply in comparison to the classical streamer mechanism [2]. When fast electrons are decelerated on an anode, x rays are emitted [3–6] and initiate secondary electrons and new avalanches. Since these processes are short-term and overlap in time, it is very difficult to analyze the dynamics of the initiation of subnanosecond gas breakdown. It is also worth noting that the regime of the continuous acceleration of electrons was analyzed primarily theoretically, because the corresponding experiments are complicated. The existence of fast electrons has been experimentally proved in only a few works [3–6]. Experimental data related to the effect of continuously accelerated electrons on the initiation of a subnanosecond gas discharge are virtually not available.

This work is devoted to an attempt to study the dynamics of the initiation and development of the breakdown of highly overvoltaged gas gaps at the initial stage (for time intervals shorter than 1 ns).

EXPERIMENTS

The experiments were carried out on a setup based on a small-voltage-pulse generator RADAN-303 [7], which provides for the synchronous detection of subnanosecond high-voltage pulses applied to a gas gap with electron–optical chronography of luminescence accompanying prebreakdown and breakdown processes in the gap. High-rate electron–optical chronography [8] is the only available method that allows for the determination of the domain of the origin of breakdown inside the discharge gap. The arrangement of the setup and experimental technology were described in detail in [9, 10].

Technical nitrogen was the gas investigated in all experiments. A high-voltage pulse (Fig. 1a) with half-height durations 0.5–3 ns, controlled amplitudes 70–150 kV, and voltage-increase rates 7×10^{13}–6×10^{14} V/s at the front was supplied to the gas-discharge gap. The electrodes were made of copper. Measurements were carried out only after the training of the electrode surfaces with several hundreds of high-voltage pulses.

The experiments showed that there are two mechanisms of the development of gas breakdown at the initial stage.

1. In the first run of experiments, voltage pulses with an amplitude of 70 kV, front 0.8–1 ns at levels 0.1–0.9, and an FWHM of about 2 ns were supplied to a gas gap (4.5 atm). In this case, the increasing voltage rate at the pulse front is equal to 7×10^{13} V/s. Figure 1b shows the configuration of discharge-gap electrodes and the electric field distribution in the gas gap. The electric field distribution is calculated by the SAM interactive code

Institute of Electrophysics, Ural Division, Russian Academy of Sciences, ul. Amundsena 106, Yekaterinburg, 620016 Russia
e-mail: stivan@iep.uran.ru

1028-3358/04/4912-0701$26.00

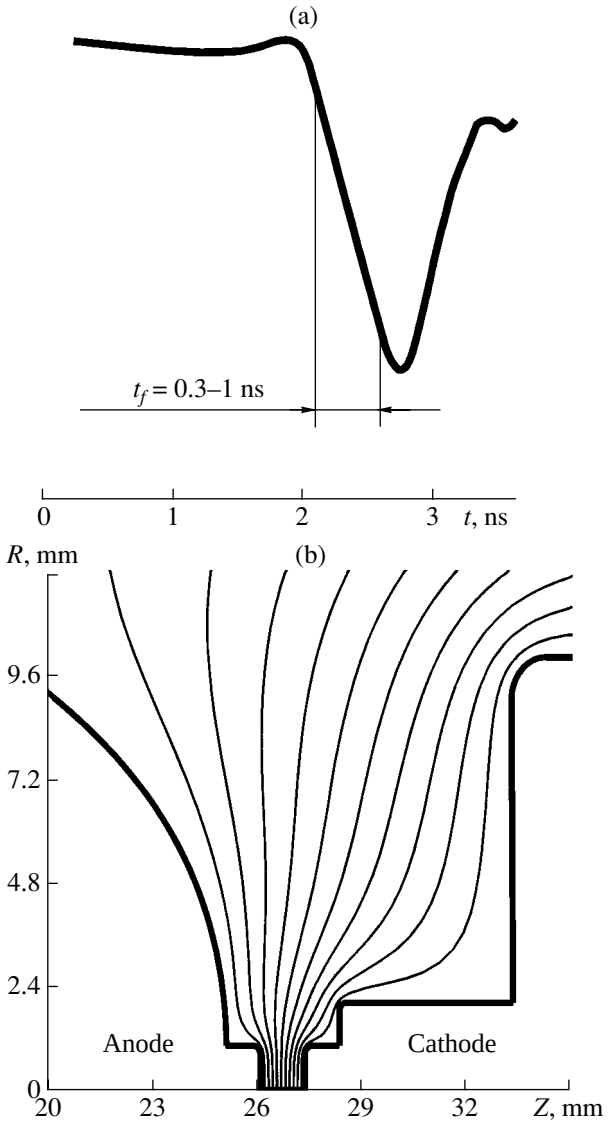


Fig. 1. (a) High-voltage pulse and (b) configuration of the electrodes and equipotential contours of the electric field in the gas gap along the Z axis of the symmetry of the cathode–anode system.

for calculating electron guns [11]. The radius of the cathode edge was taken to be 0.2 mm. The calculated value of the normal component of the electric field on the flat of the most projected part of the cathode is equal to E_1 = 285 kV/cm and E_2 = 600 kV/cm near the cathode edge for this run of experiments.

Figures 2a and 2b show typical streak photographs of luminescence accompanying breakdown. They are given in time development in the direction parallel to the electrode plane. It is seen that the process of filling the gap with luminescence occurs in three phases. In the first phase, relatively soft luminescence arises in the gas volume, begins to propagate towards the electrodes, and reaches their surfaces in 200–400 ps. In the second phase, the gap is spanned by a bright bridge, which is most pronounced in Fig. 2a. The lifetime of the bright bridge is equal to 100–200 ps. In the third phase, two processes, called phases *3a* and *3b*, are simultaneously developed. Luminescence begins to propagate from the anode surface (section *3a* of time development) and reaches the cathode in 500 ps. Simultaneously with phase *3a*, brighter luminescence arises on the surfaces of the electrodes and reaches the center of the gap in 800–900 ps (phase *3b*). This luminescence is called secondary luminescence.

It is worth noting that the luminescence accompanying the first phase of breakdown formation can arise either at a certain point of the gas gap (Fig. 2a) or in a large part of the gas-gap volume (Fig. 2b). In particular, in the photograph shown in Fig. 2a it is seen that initial luminescence arose at point *A*, which was spaced from the cathode by a distance that is 30–35% of the total length of the interelectrode gap.

Secondary bright luminescence on the electrodes of a gas diode arises on the anode simultaneously or with a 100–150-ps delay with respect to the appearance of luminescence on the cathode. However, in some photographs (several percent of the total number), it is seen that secondary luminescence can arise first only on one of the electrodes (Figs. 2c, 2d) and then on the second electrode (with a delay of several hundreds of picoseconds). In this case, the propagation rate of intense luminescence in the gas gap from the electrode where it arose with a delay is higher by a factor of 2–2.5. In one photograph, secondary intense luminescence propagates only from the cathode (Fig. 2e). In Figs. 2c–2e, the first phase of breakdown formation is not observed, and the total breakdown formation time decreases by about 400 ps.

The experiments were repeated for pressures of 4–10 atm. Voltage pulses that were supplied to the gas diode had fronts with 0.5–1-ns widths at levels 0.1–0.9 and were characterized by a maximum voltage-increase rate at the fronts of up to 2×10^{14} V/s. The electric field on the most projected part of the cathode varied in the range E_1 = 285–600 kV/cm and, near the cathode edge, E_2 = 0.6–1.1 MV/cm. Under these conditions, the total breakdown formation time decreased to 400–500 ps and the velocity of filling the gap with secondary luminescence increased to $(7–8) \times 10^8$ cm/s, but the dynamics of the developing breakdown did not qualitatively change.

2. Mechanisms of the initiation of gas breakdown at pressures of several tens of atmospheres are of special interest. As was mentioned in [12], short fronts of generated high-voltage pulses have been obtained on high-pressure gas commutators.

Experiments were carried out in an experimental chamber at nitrogen pressures of 30–40 atm. A voltage pulse with a 300–400-ps-wide front and a FWHM of 1–2 ns was supplied to the discharge gap. The voltage-increase rate at the front was equal to (4–5) ×

10^{14} V/s. The calculated value of the normal component of the electric field on the cathode plane varied within the range E_1 = 0.9–1.4 MV/cm and E_2 = 1.7–2.4 MV/cm near the cathode edge. Figure 2f shows streak photographs of luminescence accompanying gas breakdown. The development of filling the gap with luminescence was different in this case. A relatively soft luminescence that arose at the initial stage of breakdown formation in preceding experimental runs was not observed. Instead, bright luminescence quickly filled the entire gap (in a time < 100 ps). The velocity of filling of the gap with luminescence exceeds 1.4×10^9 cm/s. In this case, it is impossible to observe in detail the very beginning stage (first 100 ps) of breakdown formation in photographs, because its duration is commensurate with the time resolution of the electron–optical camera AGAT SF3M used in the experiments.

RESULTS AND DISCUSSION

First-Type Formation of Breakdown

Initial luminescence (phase *1*) arises most likely due to the ionization of a gas under the pressure of an electron avalanche propagating in it. The avalanche can be generated by both free electrons, which are always present in the gas, and electrons emitted from the cathode surface. Even 200–300 ps after the application of the voltage pulse to the gas gap, the electric field near the cathode edge exceeds 200 kV/cm. The surface even of the trained cathode includes a sufficient number of microbulges with a field gain of about 40–50 [13] that are field-emission centers. Several electron avalanches initiating breakdown can arise simultaneously. This assumption is supported by the fact that it is often impossible to determine a single point in the gas gap from which breakdown starts. Several such points can exist simultaneously. As a result, the superposition of several avalanches and the smearing of the initial section of the time development of breakdown are observed in several photographs (Fig. 2b).

An electron avalanche disturbs the electric field in the gap, which leads to the production of secondary electrons due to gas ionization in enhanced-field domains. Plasma structures, or cathode and anode streamers, begin to propagate towards the cathode and anode, respectively. The velocities of the anode and cathode streamers are equal to 5.8×10^8 and 3.1×10^8 cm/s, respectively (all calculations in this section were performed for Fig. 3a). When the anode and cathode streamers reach the electrode surfaces, a bright bridge or spark channel arises between them (phase *2*). The rate of propagation of the spark channel lies in the range $(2–4) \times 10^9$ cm/s. The formation of the spark channel induces a new ionization wave that propagates from the anode to the cathode (phase *3a*) with a velocity of 7×10^8 cm/s. Intense secondary luminescence

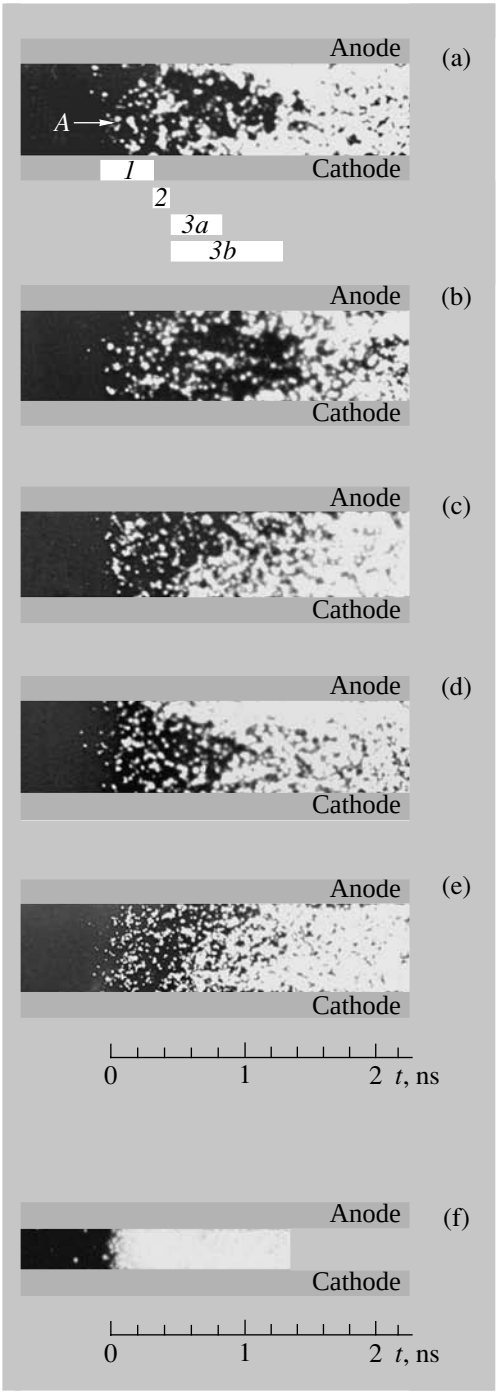


Fig. 2. Streak photographs of luminescence accompanying breakdown of gas gaps with a width of (a–e) 3.55 and (f) 1.42 mm at a pressure of (a–e) 4.5 and (f) 40 atm.

(phase *3b*) arises most likely due to the explosions of microinhomogeneities on the electrode surface. This assumption is corroborated by the observation, in some photographs, of bright luminescence that begins on only one electrode (Figs. 2c–2e); i.e., the process is random. Metal vapors produced in the explosions actively ionize the gas medium. As a result, the brightness of the

luminescence accompanying breakdown increases sharply. This luminescence propagates to the gap center with a rate of about 2×10^8 cm/s. In our experiment, the end of phase *3b* corresponds to a current of ~1.4 kA or higher in the gas gap.

The first phase of breakdown is not observed in Figs. 2c–2e. One can assume that breakdown is initiated in this case due primarily to field emission from the cathode surface and the explosion of microinhomogeneities on the electrode surface, whereas the phase of formation of a single electron avalanche (phase *1*) is absent.

Thus, we can conclude that the pulsed electrical breakdown observed in the first experimental runs was initiated by two mechanisms: ionization of the gas by electron avalanches and the explosive emission of electrons. Explosive emission processes sometimes dominate.

Second-Type Formation of Breakdown

The second type of the development of breakdown can possibly be attributed to the appearance of a large number of fast electrons in the discharge gap. Fast electrons and corresponding bremsstrahlung ionize the gas faster and more efficiently than recombination radiation or stepwise ionization. Therefore, the breakdown formation time can decrease abruptly. Owing to a high penetrability of fast electrons and bremsstrahlung, the gas far from the primary centers of ionization is also ionized. The discharge thereby loses its compact form and becomes diffuse or multichannel [14]. The criterion of the transition from the streamer mechanism of gas discharge to the continuous acceleration of electrons was formulated in [1] as $\frac{E_c}{p} = 3.88 \times 10^3 \frac{Z}{I}$, where E_c is the critical field measured in V/(cm Torr), Z is the atomic number of the gas, and I is the mean energy of inelastic losses measured in electron volts. For nitrogen, $Z = 14$, $I = 75$–80 eV, and $\frac{E_c}{p} = 590$ V/(cm Torr). The critical field E_c for nitrogen is equal to 17.3 MV/cm at 40 atm. In our experiment, the electric field near the cathode edge was equal to 1.7–2.4 MV/cm; i.e., $\frac{E_c}{E} =$ 7–10. Therefore, fast electrons must absent. However, streak photographs show that the development of breakdown is different from that observed in the first run of experiments. Moreover, $\frac{E_c}{E} = 3$ for the first run of experiments; i.e., the appearance of fast electrons was more probable. An increase in the rate of breakdown formation to 1.4×10^9 cm/s or higher, which was observed experimentally, can be explained as follows. When a plasma cloud, which is formed in explosive emission processes on the cathode and has a high conductivity, moves from the cathode to the anode, the electric field is redistributed in the gap. An electric field whose strength is higher than the critical value can appear in a certain domain of the gas gap for a very short time (apparently no more than several tens of picoseconds). This field is responsible for the production of fast electrons and the reduction of the breakdown formation time. Second-type breakdown was observed only when the front of a high-voltage pulse applied to the gas gap was narrower than 400 ps.

CONCLUSIONS

If the rate of increasing pulsed voltage applied to the discharge gap is lower than 2×10^{14} V/s and the electric field near the cathode edge lies in the range 0.6–1.1 MV/cm, the pulsed electrical breakdown observed in the experiments was initiated by two mechanisms simultaneously: ionization of the gas by electron avalanches and due to the explosive emission of electrons. Explosive emission processes sometimes dominate. When the rate of increasing voltage applied to the discharge gap is higher than 4×10^{14} V/s and the electric field lies in the range 1.7–2.4 MV/cm, the mechanism of initiating subnanosecond gas breakdown changes, and the rate of filling of the gap with luminescence increases sharply. In this case, the propagation of the discharge towards the anode is likely determined by runaway electrons. The indicated processes are of interest for engineers developing generators of subnanosecond voltage pulses. Involving fast electrons under optimum conditions can significantly increase the rate of breakdown formation, which will enable one to develop new types of superfast high-pressure gas commutators.

ACKNOWLEDGMENTS

I am grateful to Corresponding Member of the RAS V.G. Shpak, S.A. Shunaĭlov, and Corresponding Member of the RAS M.I. Yalandin for assistance in the creation of the experimental setup and useful advice.

REFERENCES

1. Yu. D. Korolev and G. A. Mesyats, *Physics of Pulsed Breakdown in Gas* (Nauka, Moscow, 1991; Ural Division of the RAS, Yekaterinburg, 1998).
2. H. Raether, *Electron Avalanches and Breakdown in Gases* (Butterworths, London, 1964; Mir, Moscow, 1968].
3. L. V. Tarasova, L. N. Khudyakova, T. V. Loĭko, *et al.*, Zh. Tekh. Fiz. **44**, 564 (1974) [Sov. Phys. Tech. Phys. **44**, 351 (1974)].
4. R. C. Noggle, E. P. Kriger, and J. R. Wayland, J. Appl. Phys. **39**, 4746 (1968).
5. Yu. L. Stankevich and N. S. Kalinin, Dokl. Akad. Nauk SSSR **177** (1), 72 (1967) [Sov. Phys. Dokl. **12**, 1042 (1967)].

6. V. V. Kremnev and Yu. A. Kurbatov, Zh. Tekh. Fiz. **42**, 795 (1972) [Sov. Phys. Tech. Phys. **42**, 626 (1972)].
7. V. G. Shpak, S. A. Shunaĭlov, M. I. Yalandin, *et al.*, Prib. Tekh. Eksp., No. 1, 149 (1993).
8. M. M. Butslov, B. M. Stepanov, and S. D. Fanchenko, *Electrical-to-Optical Converters and Their Application in Scientific Research* (Nauka, Moscow, 1978) [in Russian].
9. S. N. Ivanov, V. G. Shpak, S. A. Shunaĭlov, *et al.*, Prib. Tekh. Eksp., No. 5, 51 (2000).
10. S. N. Ivanov, V. G. Shpak, S. A. Shunailov, *et al.*, in *Proceedings of the 13th International Symposium on Gas Discharges and Their Applications, Glasgow, 2000*, pp. 497–500.
11. M. A. Tiunov, B. M. Fomel', and V. P. Yakovlev, Preprint No. 89, IYaF SO AN SSSR (Institute of Nuclear Physics, Siberian Division, Academy of Sciences of the USSR, Novosibirsk, 1989).
12. G. A. Mesyats, *Ectons. Part 3. Ectons in Electrophysical Devices* (UIF Nauka, Yekaterinburg, 1994) [in Russian].
13. M. I. Elinson and G. F. Vasil'ev, *Field Emission* (Fizmatgiz, Moscow, 1958) [in Russian].
14. Yu. L. Stankevich, Zh. Tekh. Fiz. **40**, 1476 (1970) [Sov. Phys. Tech. Phys. **40**, 1138 (1970)].

Translated by R. Tyapaev

Doklady Physics, Vol. 49, No. 12, 2004, pp. 706–709. Translated from Doklady Akademii Nauk, Vol. 399, No. 4, 2004, pp. 477–480.

PHYSICS

Mechanism of Stochastic Resonance

P. S. Landa

Presented by Academician A.M. Dykhne April 19, 2004

Received April 20, 2004

Stochastic resonance in an overdamped oscillator is considered theoretically. It has been shown that a seeming resonance is actually caused by a noise-induced change in the effective stiffness and damping factor with respect to a signal. For a certain noise intensity, the effective stiffness is minimal, which leads to a nonmonotonic variation of the output-signal amplitude as a function of noise intensity. It is substantial that the position of the minimum of the effective stiffness and its value depend strongly on the signal frequency. The results are compared with similar processes for vibrational resonance. Considerable differences between these phenomena are indicated.

For the last 25 years, numerous phenomena of different origins but that are referred to as resonances have been explored in many works. Any resonance-like nonmonotonic dependence of a variable in the system under consideration on the intensity of an external action or on another variable was treated as resonance. In most cases, this resonance is not similar to the classical resonance that arises when some frequencies of the system coincide with each other or are divisible. Nevertheless, these phenomena are frequently observed and are interesting both theoretically and practically. It is worth noting that several authors have mentioned that the term "stochastic resonance" is inadequate (see, e.g., [1–4]).

At present, two types of resonances have been distinguished: stochastic resonance, where the maximum of a certain variable is reached by varying the intensity (spectral density) of noise [5, 6], and vibrational resonance, where a similar effect is reached by varying the amplitude of a high-frequency harmonic external force [4]. Most theoretical and experimental works are devoted to stochastic resonance. The first work on this subject appeared in 1981 primarily in connection with the necessity of explaining the alternation of Earth's ice ages, a cycle that is close to periodic (with a period of about 100 000 yr) [7, 8]. Vibrational resonance was discovered much later [9]. However, since it is simpler than stochastic resonance, its rigorous theory has been developed, and a new type of bifurcations, namely, vibration-induced bifurcations, has been discovered [10, 11].

Stochastic resonance is most often considered in terms of an example involving the following equation of motion of a light particle in a bistable potential field perturbed by a weak periodic signal and noise:

$$\dot{x} + f(x) = A\cos\omega t + \xi(t), \qquad (1)$$

where x is the displacement of the particle, $A\cos\omega t$ is the weak periodic perturbation with frequency ω, $f(x) = \frac{dU(x)}{dx}$ is the force acting on the particle, $U(x)$ is the symmetric double-well potential (the simplest potential $U(x) = -\frac{x^2}{2} + \frac{x^4}{4}$ will be considered), and $\xi(t)$ is white noise with intensity K; i.e., $\langle\xi(t)\xi(t+\tau)\rangle = K\delta(\tau)$.

According to [12], the spectrum of Eq. (1) contains discrete components (odd harmonics of frequency ω) and continuous components induced by noise. The ratio of the amplitude B of the first harmonic of the signal at the output of the system to the signal amplitude A at its input is called the gain factor $Q(K)$.

Numerical experiments show that the gain factor $Q(K)$ and phase shift $\psi(K)$ between the output and input signals are nonmonotonic functions of the noise intensity K [4]. The gain factor Q reaches a maximum for a certain K value, which increases with the signal frequency. For a fixed frequency and amplitude of the signal, Q as a function of K is similar to the resonance dependence of the amplitude of an oscillator on the frequency of the driving force. Since K determines the mean frequency of noise-induced jumps from one well to the other, which are associated with random transitions through the potential barrier, it was first assumed that the maximum in the K dependence of Q is reached when the signal period $T = \frac{2\pi}{\omega}$ is equal to double the mean time of the first transition through the potential barrier. This assumption implies that Q has a maximum not only as a function of K but also as a function of ω.

Moscow State University,
Vorob'evy gory, Moscow, 119992 Russia
e-mail: planda@lpi.ru

1028-3358/04/4912-0706$26.00

However, it is known that Q is a monotonically decreasing function of ω [4]. Therefore, the above assumption is obviously invalid.

The true cause of the resonance K dependence of Q is a nonmonotonic noise-induced variation of the effective parameters of the system [4]. The possibility of the noise-induced variation of the effective parameters of the system has long been known (see, e.g., [13]). However, stochastic resonance was previously considered from this point of view only in my works [4]. In order to calculate the effective parameters, the solution of Eq. (1) is represented in the form

$$x(t) = s(t) + n(t), \tag{2}$$

where $s(t) = \langle x(t)\rangle$, $\langle n(t)\rangle = 0$, and this equation is decomposed into two equations. One of these equations describes quantities averaged over the statistical ensemble, and the other equation, deviations from the mean values. Further calculations are performed under the assumption that the input-signal amplitude A is low. Since all odd moments of noise $m_j = \langle n^j\rangle$ are equal to zero in the absence of the signal s, we can set $m_3 = as + b\dot{s}$ in the linear approximation, where a and b are unknown coefficients that will be determined below. In this approximation, the equations for $s(t)$ and $n(t)$ have the form

$$(1 + b_1)\dot{s} + ks + s^3 = A\cos\omega t, \tag{3}$$

$$\dot{n} + (3s^2 - 1)n + n^3 + (3n^2 - 1 - k)s - b_1\dot{s} = \xi(t), \tag{4}$$

where

$$k = 3m_2 - 1 + a_1 \tag{5}$$

is the effective stiffness of the system with respect to the signal $s(t)$. The parameters a_1 and b_1 are introduced in addition to a and b in order to satisfy Eqs. (10). As was expected, calculations showed that $a_1 = a$ and $b_1 = b$.

It is worth noting that such a separation of equations is similar to the separation of motion into fast and slow components that was proposed by Blekhman [14]. However, in the case under consideration, the motion is separated into regular and random rather than into fast and slow.

According to Eq. (3), the variable $s(t)$ can oscillate either near zero if $3m_2 - 1 + a > 0$ or near one of the equilibrium positions $s_0 = \pm(1 - 3m_2 - a)$ if $3m_2 - 1 + a < 0$. It can be shown that the second case is impossible.

To calculate the moments m_j and to obtain a, b, a_1, and b_1, we use the Fokker–Planck equation corresponding to Langevin equation (4). In the linear approximation in $s(t)$, it has the form

$$\frac{\partial w}{\partial t} = \frac{\partial}{\partial n}\{[n^3 - n + (3n^2 - 1 - k)s - b_1\dot{s}]w\} + \frac{K}{2}\frac{\partial^2 w}{\partial n^2}. \tag{6}$$

It is convenient to separate the solution of Eq. (6) in the steady state into the following three components:

$$w(n, t) = w_0(n) + w_1(n)s(t) + w_2(n)\dot{s}(t). \tag{7}$$

Since $s(t)$ in the linear approximation is a harmonic signal of frequency ω, $\ddot{s}(t) = -\omega^2 s(t)$. Therefore, substituting Eq. (7) into Eq. (6) and equating the corresponding terms, we arrive at the following equations for these components:

$$\frac{K}{2}\frac{\partial w_0}{\partial n} + (n^3 - n)w_0 = 0,$$
$$w_1 = \frac{d}{dn}((n^3 - n)w_2) + \frac{K}{2}\frac{d^2 w_2}{dn^2} - b_1\frac{dw_0}{dn}, \tag{8}$$

$$-\omega^2 w_2 = \frac{d}{dn}((n^3 - n)w_1) + \frac{K}{2}\frac{d^2 w_1}{dn^2} + \frac{d}{dn}((3n^2 - 1 - k)w_0). \tag{9}$$

We note that $w_0(n)$ is an even function of n, whereas $w_1(n)$ and $w_2(n)$ are odd functions of n. Therefore,

$$m_2 = \int_{-\infty}^{\infty} n^2 w_0(n)dn,$$
$$\int_{-\infty}^{\infty} nw_1(n)dn = 0, \quad \int_{-\infty}^{\infty} nw_2(n)dn = 0, \tag{10}$$
$$a = \int_{-\infty}^{\infty} n^3 w_1(n)dn, \quad b = \int_{-\infty}^{\infty} n^3 w_2(n)dn.$$

Solving Eq. (8), we obtain

$$w_0(n) = C\exp\left(\frac{n^2}{K} - \frac{n^4}{2K}\right), \tag{11}$$

where C is the normalization constant.

It is convenient to represent the solution of Eqs. (9) in the form

$$w_1(n) = w_{11}(n) + kw_{12}(n) + b_1 w_{13}(n),$$
$$w_2(n) = w_{21}(n) + kw_{22}(n) + b_1 w_{23}(n), \tag{12}$$

where the functions $w_{11}(n)$, $w_{12}(n)$, $w_{13}(n)$, $w_{21}(n)$,

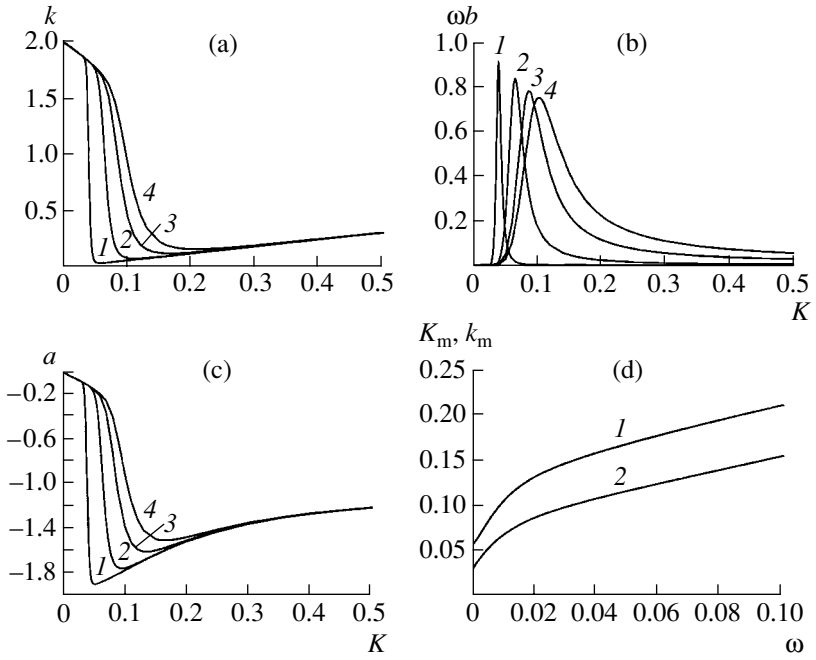


Fig. 1. (a) k, (b) ωb, and (c) a vs. noise intensity K for ω = (*1*) 0.0001, (*2*) 0.01, (*3*) 0.05, and (*4*) 0.1 and (d) signal-frequency dependence of (*1*) K_m and (*2*) k_m.

$w_{22}(n)$, and $w_{23}(n)$ satisfy the equations

$$w_{11} = \frac{d}{dn}((n^3 - n)w_{21}) + \frac{K}{2}\frac{d^2 w_{21}}{dn^2},$$

$$-\omega^2 w_{21} = \frac{d}{dn}((n^3 - n)w_{11})$$

$$+ \frac{K}{2}\frac{d^2 w_{11}}{dn^2} + \frac{d}{dn}((3n^2 - 1)w_0(n)),$$

$$w_{12} = \frac{d}{dn}((n^3 - n)w_{22}) + \frac{K}{2}\frac{d^2 w_{22}}{dn^2}, \qquad (13)$$

$$-\omega^2 w_{22} = \frac{d}{dn}((n^3 - n)w_{12}) + \frac{K}{2}\frac{d^2 w_{12}}{dn^2} - \frac{dw_0(n)}{dn},$$

$$w_{13} = \frac{d}{dn}((n^3 - n)w_{23}) + \frac{K}{2}\frac{d^2 w_{23}}{dn^2} - \frac{dw_0(n)}{dn},$$

$$-\omega^2 w_{23} = \frac{d}{dn}((n^3 - n)w_{13}) + \frac{K}{2}\frac{d^2 w_{13}}{dn^2}.$$

Equations (13) must be solved with zero boundary conditions for $n = \pm\infty$. Moreover, the functions $w_{11}(n)$, $w_{12}(n)$, $w_{13}(n)$, $w_{21}(n)$, $w_{22}(n)$, and $w_{23}(n)$ must vanish for $n = 0$, because they are odd functions.

From Eqs. (10) and (12), the equations for the unknowns a, b, k, and b_1 follow in the form

$$J_{11} + kJ_{12} + b_1 J_{13} = 0, \quad J_{21} + kJ_{22} + b_1 J_{23} = 0,$$
$$a = I_{11} + kI_{12} + b_1 I_{13}, \quad b = I_{21} + kI_{22} + b_1 I_{23}, \qquad (14)$$

where

$$J_{ij} = \int_0^\infty n w_{ij}(n)dn, \quad I_{ij} = \int_{-\infty}^\infty n^3 w_{ij}(n)dn.$$

Solving these equations and taking Eq. (5) into account, we obtain

$$k = \frac{J_{13}J_{21} - J_{23}J_{11}}{J_{12}J_{23} - J_{22}J_{13}}, \quad b_1 = \frac{J_{11}J_{22} - J_{12}J_{21}}{J_{12}J_{23} - J_{22}J_{13}},$$

$$a_1 = k - 3m_2 + 1, \qquad (15)$$

$$a = I_{13}b_1 + I_{12}k + I_{11}, \quad b = I_{23}b_1 + I_{22}k + I_{21}.$$

According to the above results, the effective stiffness k and correction to the damping factor b depend strongly on the signal frequency ω. Figure 1 shows k, ωb, and a as functions of noise intensity K for some frequencies ω. As is seen, the value $K = K_m$ is close to zero for low ω values and increases with ω (Fig. 1d).

Taking into account that the imaginary part of the effective stiffness c_i calculated numerically in [4] is equal to ωb, we conclude that the dependences obtained above qualitatively correspond to those presented in [4], although there are quantitative differences.

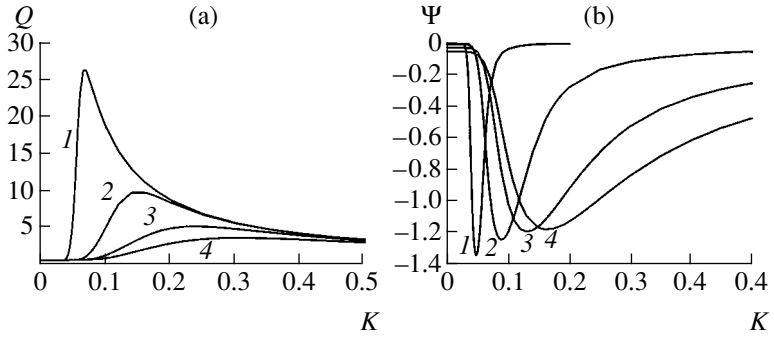


Fig. 2. (a) Q and (b) ψ vs. noise intensity K for ω = (*1*) 0.0001, (*2*) 0.01, (*3*) 0.05, and (*4*) 0.1.

In the case under consideration, K_m has a meaning similar to the bifurcation amplitude of high-frequency vibration for vibrational resonance, because the existence of this minimum is responsible for the observed resonance-like dependence of the output signal amplitude on noise intensity. However, since this minimum is nonzero, it does not lead to the bifurcation of the change in the number of equilibrium states. This circumstance is one of the significant differences between stochastic resonance and vibrational resonance. The second significant difference is that the correction to the damping factor is absent for vibrational resonance in the approximation under consideration, whereas this correction is large for stochastic resonance.

With known k and b values, it is easy to calculate the gain factor and phase by the formulas

$$Q = \frac{B}{A} = \sqrt{\frac{1}{\omega^2(1+b)^2 + k^2}}, \qquad \psi = -\arctan\frac{\omega(1+b)}{k}. \tag{16}$$

Figure 2 shows Q and ψ as functions of K for ω = 0.0001, 0.01, 0.05, and 0.1. It is worth noting that the resonance value of noise intensity for each ω value is much higher than $K_m(\omega)$. This is due to the correction to the damping factor.

REFERENCES

1. R. F. Fox, Phys. Rev. A **39**, 4148 (1989).
2. M. I. Dykman *et al.*, J. Stat. Phys. **70**, 479 (1993).
3. Yu. L. Klimontovich, Usp. Fiz. Nauk **169**, 39 (1999) [Phys. Usp. **42**, 37 (1999)].
4. P. S. Landa, *Regular and Chaotic Oscillations* (Springer, Heidelberg, 2001).
5. L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998).
6. V. S. Anishchenko, A. B. Neĭman, F. Moss, *et al.*, Usp. Fiz. Nauk **169**, 7 (1999) [Phys. Usp. **42**, 7 (1999)].
7. R. Benzi, A. Sutera, and A. Vulpiani, J. Phys. A **14**, L453 (1981).
8. G. Nicolis and C. Nicolis, Tellus **33**, 225 (1981).
9. P. Landa and P. V. E. McClintock, J. Phys. A **33**, L433 (2000).
10. I. I. Blekhman and P. S. Landa, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelin. Dinamika **10**, 44 (2002).
11. I. I. Blekhman and P. S. Landa, Int. J. Non-Linear Mech. **39**, 421 (2004).
12. R. L. Stratonovich, *Selected Problems of the Theory of Fluctuations in Radio Engineering* (Sovetskoe Radio, Moscow, 1961) [in Russian].
13. A. A. Pervozvanskiĭ, *Random Processes in Nonlinear Automation Systems* (Fizmatgiz, Moscow, 1962) [in Russian].
14. I. I. Blekhman, *Vibrational Mechanics* (Nauka, Moscow, 1994) [in Russian].

Translated by R. Tyapaev

Doklady Physics, Vol. 49, No. 12, 2004, pp. 710–714. Translated from Doklady Akademii Nauk, Vol. 399, No. 4, 2004, pp. 481–485.

PHYSICS

Features of the Excitation of $^1F^\circ$ Series of the Ytterbium Atom by Slow Electrons

Yu. M. Smirnov

Presented by Academician A.I. Leont'ev May 28, 2004

Received May 28, 2004

1. Only two decades ago, ytterbium was considered to be an uninteresting element "about which almost nothing can be told" [1]. Since that time, the situation has changed, and ytterbium should undoubtedly be referred to as an important element of modern quantum electronics.

Coherent optical radiation was generated via transitions in both the ytterbium atom and the single-charged ytterbium ion [2, 3]. More recently, high-temperature superconductivity was obtained in an $YbBa_2Cu_3O_x$ thin film applied by molecular beam epitaxy [4]. In addition, a very high content of ytterbium (and other rare-earth elements) was found in the atmospheres of rapidly oscillating stars with anomalous chemical compositions (*roAp* stars), which have recently been actively investigated [5, 6]. Thus, during the last two decades, ytterbium, which was once considered to be an uninteresting element, has instead become a very interesting element.

At the same time, investigations of the properties of ytterbium in the gas phase (primarily the determination of its atomic constants) are lagging far behind advances in its application. The radiation constants of the ytterbium atom are more well-determined: the probabilities of transitions for 52 spectral lines of YbI were experimentally determined in [7], and lifetimes for 62 energy levels of YbI were presented in review [8].

The collision constants of the ytterbium atom are less well-known. The electron-impact excitation of spectral lines of YbI was experimentally studied in [9], where excitation cross sections for 34 spectral lines of YbI for electron energies of 100, 200, and 300 eV, as well as 15 optical excitation functions (OEFs), were presented. For all the lines under investigation, the cross sections were determined at the maximums of the OEFs. The cross sections for the excitation of the ytterbium atom were previously calculated in [10] in the Born and Ochkur approximations. However, the author of that paper presented the calculated parameters α, β, and γ, which appear in the asymptotic relation for calculating cross sections, rather than the cross sections themselves. In the same paper, it was shown by an example of the excitation of two *P* levels of the europium atom that the calculated cross sections are two orders of magnitude larger than the experimental data. Almost the same discrepancy exists for the ytterbium atom.

The excitation of transitions in the double-charged ytterbium ion (excitation with simultaneous double ionization) was experimentally studied in more recent work [11]. The difference between the theoretical and experimental cross sections was analyzed in the same work. It was shown that the cross sections obtained in [11] agree well with theoretical estimates in order of magnitude and differ from the experimental data [9] by one or two orders of magnitude.

It is worth noting that excitation of only the *S*, *P*, and *D* levels of the ytterbium atom was studied in [9, 10]; the excitation of states with higher orbital angular momenta was not considered. The same is true for the lifetimes of YbI presented in [8]. In this work, the excitation of $^1F_3^\circ$ levels of the ytterbium atom, as well as two transitions from the $4f^{13}5d6s6p\ ^1H_5$ level and one transition from the $4f^{13}6s^26p\ ^1F_3$ level, are analyzed for the first time. The parity of the two last levels coincides with the parity of the ground state of the ytterbium atom, $4f^{14}(^1S)6s^2\ ^1S_0$, which is the initial level in the excitation process.

2. The experiment was carried out by the method of extended crossing beams. The basic concept of this method and its technical realization were discussed in detail in [12]. Therefore, it is unnecessary to discuss them here. It is only necessary to present the basic experimental parameters that are specific to ytterbium.

To create an atomic beam, ytterbium was evaporated from a tantalum crucible heated by an electron beam up to 800 K. Since ytterbium is an easily evaporated metal, the atomic density in the overlapping domain of the

Moscow Power Engineering Institute (Technical University), ul. Krasnokazarmennaya 14, Moscow, 111250 Russia
e-mail: smirnovYM@mpei.ru

1028-3358/04/4912-0710$26.00

Excitation cross section for the ytterbium atom

λ, nm	Transition	J	E_{low}, cm^{-1}	E_{up}, cm^{-1}	Q_{50}, 10^{-18} cm^2	Q_{max}, 10^{-18} cm^2	$E(Q_{max})$, eV	OEF
428.210	$4f^{14}5d6s\ ^3D–4f^{14}6s8f\ ^1F^\circ$	2–3	24751	48098	0.85	1.10	18	5
437.945	$4f^{14}5d6s\ ^3D–4f^{14}6s8f\ ^1F^\circ$	3–3	25270	48098	5.02	6.52	18	5
454.687	$4f^{14}5d6s\ ^3D–4f^{14}6s7f\ ^1F^\circ$	3–3	25270	47257	0.59	0.74	25	3
459.083	$4f^{13}5d6s^2\ ^3H^\circ–4f^{13}5d6s6p\ ^1H$	5–5	25859	47636	0.47	0.59	24	4
473.652	$4f^{14}5d6s\ ^3D–4f^{14}6s6f\ ^1F^\circ$	2–3	24751	45858	0.14	0.22	24	2
489.561	$4f^{14}5d6s\ ^1D–4f^{14}6s8f\ ^1F^\circ$	2–3	27677	48098	2.60	3.37	18	5
491.959	$4f^{13}5d6s^2\ ^3H^\circ–4f^{13}5d6s6p\ ^1H$	6–5	27314	47636	0.040	0.050	24	4
510.575	$4f^{14}5d6s\ ^1D–4f^{14}6s7f\ ^1F^\circ$	2–3	27677	47257	9.05	11.3	25	3
539.062	$4f^{14}5d6s\ ^3D–4f^{14}6s5f\ ^1F^\circ$	2–3	24751	43297	4.81	8.47	17	1
549.875	$4f^{14}5d6s\ ^1D–4f^{14}6s6f\ ^1F^\circ$	2–3	27677	45858	18.7	28.8	24	2
554.581	$4f^{14}5d6s\ ^3D–4f^{14}6s5f\ ^1F^\circ$	3–3	25270	43297	0.43	0.75	17	1
640.035	$4f^{14}5d6s\ ^1D–4f^{14}6s5f\ ^1F^\circ$	2–3	27677	43297	69.8	123.0	17	1
792.240	$4f^{13}5d6s^2\ ^3P^\circ–4f^{13}6s^26p\ ^1F$	2–3	23188	35807	2.11	–	–	–

atomic and electron beams reached 2.8×10^{10} cm^{-3} even at such a low temperature. However, any effect of the reabsorption of spectral lines for such a density was excluded, because none of the upper levels whose excitations were investigated in this work combines with the ground state of YbI.

The ground state of the ytterbium atom, $4f^{14}(^1S)6s^2\ ^1S_0$, is an isolated level spaced from the nearest excited $4f^{14}(^1S)6s6p\ ^3P_0^\circ$ level by an energy gap wider than 17000 cm^{-1}. For the low evaporation temperature indicated above, the population of any excited levels due to thermal excitation is negligibly low. Therefore, the ground state is the only initial level for the excitation of ytterbium in this experiment.

The current density of the electron beam in the working energy range 0–200 eV was less than 1.0 mA/cm^2. The width of the energy distribution of beam electrons in this energy range did not exceed 1.0 eV (for 90% of electrons). The actual spectral resolution of the optical system, including the time characteristics of the signal-detection system, was equal to about 0.1 nm. Since the YbI spectral lines considered in this work are not among the intense spectral lines of the ytterbium atom, it was impossible to increase resolution by a further decrease in the width of the slits of the monochromator.

3. The 199–847-nm emission optical spectrum, which arises due to collisions between ytterbium atoms and 50-eV electrons, was detected. About 800 spectral lines belonging to the YbI, YbII, and YbIII spectra were identified on spectrograms. For most of the lines of these three objects, OEFs were measured for electron energies in the range 0–200 eV. In this work, we discuss the excitation of spontaneous transitions from ytterbium's atomic levels with orbital angular momenta $L \geq 3$, which have not yet been studied.

The results are given in the table, along with necessary additional information on the characteristics of the transitions. The table presents the wavelength λ; the transition and internal quantum number J; the energies E_{low} and E_{up} of the lower and upper levels, respectively; the cross sections Q_{50} and Q_{max} for an electron energy of 50 eV and at the maximum of the OEF, respectively; and the position $E(Q_{max})$ of the maximum. The numbers in the OEF column correspond to the numbers of the OEFs drawn in Fig. 1. As was accepted for the representation of atomic OEFs, the abscissa scale is logarithmic and the ordinate scale is linear. In order to avoid the crossing or touching of lines, each OEF is drawn from its individual zero. Each OEF is normalized to unity at the maximum.

As is seen in the table, the data on transitions from even levels are scanty, and it is premature to discuss them. At the same time, the data on the excitation of $4f^{14}6snf\ ^1F_3^\circ$ levels are more systematic and enable one to reveal some excitation features of these levels. It

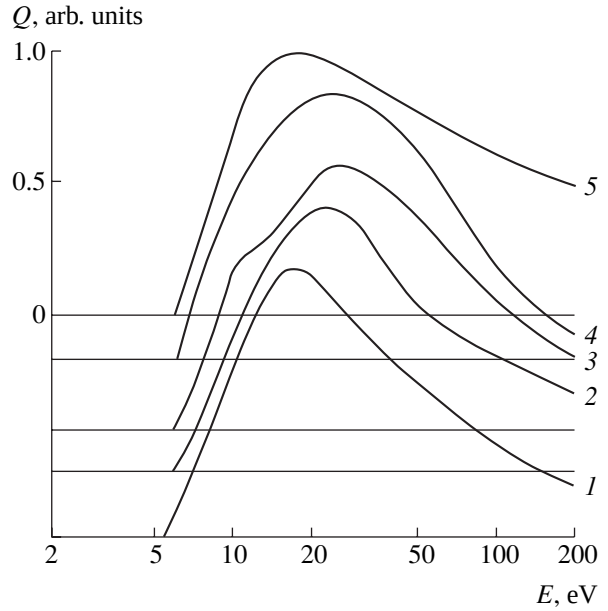


Fig. 1. Optical excitation functions for the ytterbium atom.

should be noted that all $^1F^\circ$ levels being investigated are excited due to optically forbidden transitions, because they are excited from the atomic ground state through the $6s \longrightarrow nf$ transition of one of equivalent *s* electrons. Thus, the measured cross sections are not as large as for resonance transitions and do not exceed 10^{-16} cm^2 for any line.

Figure 2 shows the $^1F^\circ$ levels of YbI with the transitions being studied. Only four low levels with n = 5–8 are presented, and these are all the available data on the positions of the $^1F^\circ$ levels of the ytterbium atom. According to [13], the positions of only these $^1F^\circ$ levels of YbI have been determined; data on the $^1F^\circ$ levels are absent in all other works. The position of the $4f^{14}6s9f\,^1F_3^\circ$ level was also presented in [13]. However, it was given as presumable, and no transition from this level was indicated. In this work, such transitions were also not observed.

As is seen in the state diagram, branching is observed for all levels under investigation. In addition to the completely allowed $4f^{14}5d6s\,^1D_2$–$4f^{14}6snf\,^1F_3^\circ$ transitions, intercombination transitions to the triplet $4f^{14}5d6s\,^3D_{2,3}$ levels occur. According to the table, excitation cross sections for intercombination transitions for levels with n = 5–7 are one or two orders of magnitude smaller than those for allowed transitions (those occurring without changes in the multiple order). Two intercombination transitions were not detected, because their excitation cross sections were below the sensitivity of the instrument. This result is not surprising,

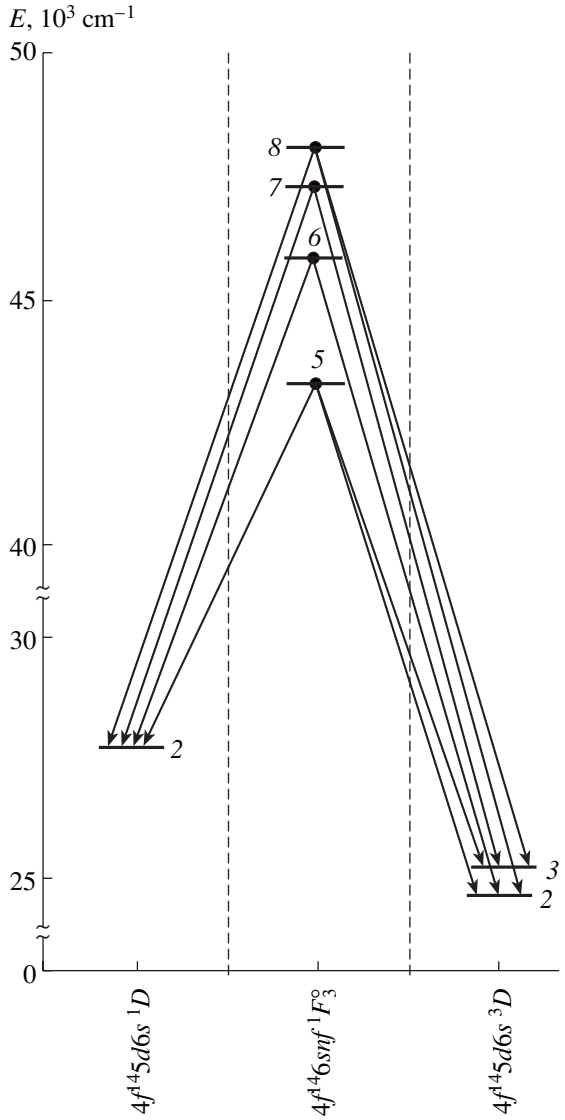


Fig. 2. Partial state diagram for YbI with the transitions under investigation.

because a similar behavior of the intensities of spectral lines was observed in [13].

As is known (see, e.g., [14]), the excitation cross sections Q for unperturbed spectral series are power functions of the principal quantum number of the upper level n:

$$Q = A_i n^{-\alpha_i}, \qquad (1)$$

where the constants A_i and α_i are unique for each spectral series. Perturbation can be manifested through three main effects [14]: first, the deviation of dependence $Q = f(n)$ from power law (1); second, a significant change in the shape of the OEF; and, third, a sharp change in the branching character.

Figure 3 shows the dependence $Q = f(n)$ drawn by using the data taken from the table for three spectral series of YbI. Power law (1) is represented by a straight line in the double logarithmic scale. The slope of this straight line to the abscissa axis determines α_i, and the extrapolation of this straight line to $n = 1$ gives A_i. According to Fig. 3, the dependence $Q = f(n)$ for any of the three spectral series of YbI under consideration is not an exact power law. Small deviations of the line from the power law exist even for the most intense allowed $4f^{14}5d6s\,^1D_2$–$4f^{14}6snf\,^1F_3^\circ$ series. The $Q = f(n)$ dependences for two intercombination series are very far from the power law and are strongly nonmonotonic. The cross section for one of intercombination transitions for the $n = 8$ level is larger than that for the competitive resonance transition. Moreover, it is seen in Fig. 1 that the shapes of OEFs 1–3 and 5 (OEF 4 does not correspond to $^1F^\circ$ levels) differ significantly from each other. As was mentioned above, this difference is a manifestation of the perturbation of levels.

Detailed data on the perturbation of the ytterbium-atom levels under consideration are not available. Nevertheless, it was mentioned in [13] that mixing of configurations takes place for the $^3F_3^\circ$ and $^1F_3^\circ$ levels. However, this statement apparently concerns primarily triplet levels. For the levels under consideration, it was only indicated that the basic component amounts to 63 and 95% for the $4f^{14}6s5f\,^1F_3^\circ$ and $4f^{14}6s6f\,^1F_3^\circ$ levels, respectively; i.e., the latter level is almost pure. Neither the content of the leading admixture nor its characteristics were indicated. There are no data on mixing for higher levels with $n = 7$ and 8. Nevertheless, according to our data as presented in Figs. 1 and 3, mixing plays a noticeable role for the $^1F^\circ$ levels of the ytterbium atom in the range $n = 5$–8.

It should be noted that the perturbation of levels is manifested in the behavior of not only excitation cross sections, but also of the radiative lifetimes of the levels. The effect of perturbation on the lifetimes of levels, including ytterbium atomic levels, was analyzed theoretically [15]. Unfortunately, a similar analysis of excitation cross sections for ytterbium or another atom is absent.

4. Thus, the excitation cross sections for transitions in three spectral series of YbI, which is caused by the population of the $^1F^\circ$ levels that is due to collisions between electrons and ytterbium atoms, were studied. Perturbation was found to lead to a comparatively small deviation of the dependence $Q = f(n)$ from the power law for the allowed $4f^{14}5d6s\ ^1D_2$–$4f^{14}6snf\ ^1F_3^\circ$ series, whereas its effect on the shape of the OEF and branching character is strong. The absence of data on the $^1F^\circ$ levels of YbI with $n \geq 9$ is indirect evidence of the significant perturbation of these levels.

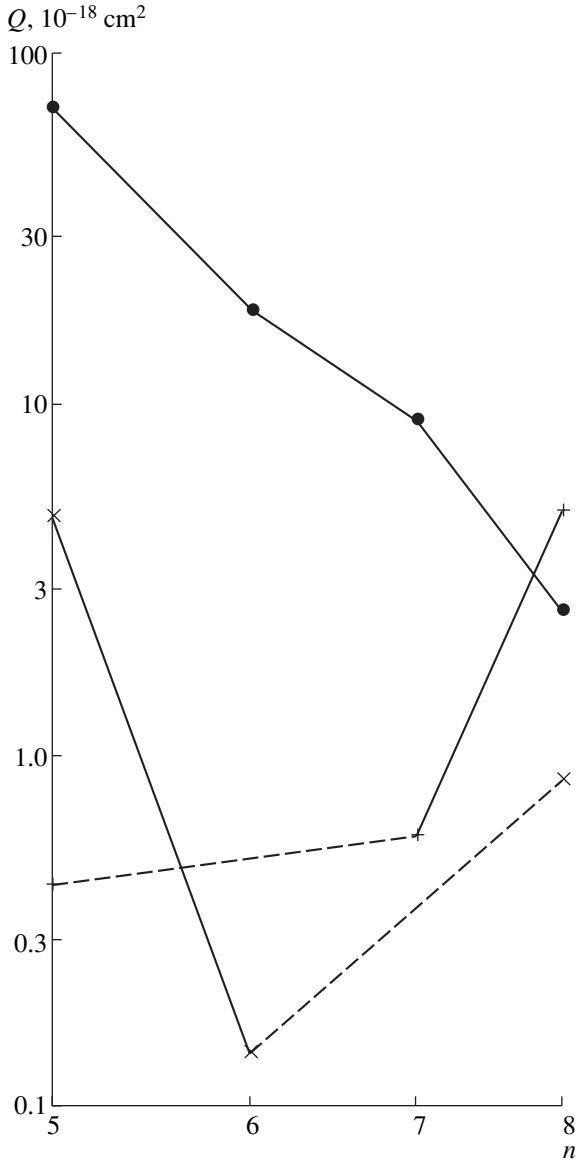


Fig. 3. Dependence $Q = f(n)$ for the (●) $4f^{14}5d6s\,^1D_2$–$4f^{14}6snf\,^1F_3^\circ$, (×) $4f^{14}5d6s\,^3D_2$–$4f^{14}6snf\,^1F_3^\circ$, and (+) $4f^{14}5d6s\,^3D_3$–$4f^{14}6snf\,^1F_3^\circ$ spectral series of the ytterbium atom.

REFERENCES

1. *Popular Library of Chemical Elements*, Ed. by I. V. Petryanov-Sokolov (Nauka, Moscow, 1983), Vol. 2, p. 157 [in Russian].
2. Ph. Cahuzac, Phys. Lett. A **27A**, 473 (1968).
3. V. M. Klimkin, Kvant. Élektron. (Moscow) **2** (3), 579 (1975) [Sov. J. Quantum Electron. **2** (3), 579 (1975)].

4. A. I. Golovashkin and E. V. Pechen', Preprint No. 98, FIAN (Lebedev Physical Institute, Academy of Sciences of the USSR, Moscow, 1989).

5. T. A. Ryabchikova, I. S. Savanov, A. P. Hatzes, *et al.*, Astron. Astrophys. **357**, 981 (2000).

6. C. R. Cowley, T. A. Ryabchikova, F. Kupka, *et al.*, Mon. Not. R. Astron. Soc. **317**, 299 (2000).

7. C. H. Corliss and W. R. Bozeman, *Experimental Transition Probabilities for Spectral Lines of Seventy Elements* (US Government Printing Office, Washington, D.C., 1962; Mir, Moscow, 1968).

8. K. B. Blagoev and V. A. Komarovskii, At. Data Nucl. Data Tables **56**, 1 (1994).

9. L. L. Shimon, N. V. Golovchak, I. I. Garga, and I. V. Kurta, Opt. Spektrosk. **50**, 1037 (1981) [Opt. Spectrosc. **50**, 571 (1981)].

10. R. K. Peterkop, Opt. Spektrosk. **48**, 10 (1980) [Opt. Spectrosc. **48**, 4 (1980)].

11. Yu. M. Smirnov, Zh. Prikl. Spektrosk. **63**, 535 (1996).

12. Yu. M. Smirnov, J. Phys. (France) **4**, 23 (1994).

13. J.-F. Wyart and P. Camus, Phys. Scr. **20**, 43 (1979).

14. Yu. M. Smirnov, Opt. Spektrosk. **92**, 400 (2002) [Opt. Spectrosc. **92**, 357 (2002)].

15. X.-W. Liu and Z.-W. Wang, Phys. Rev. A **40**, 1838 (1989).

Translated by R. Tyapaev

Doklady Physics, Vol. 49, No. 12, 2004, pp. 715–722. Translated from Doklady Akademii Nauk, Vol. 399, No. 5, 2004, pp. 609–616.

PHYSICS

Tetragonalization of $(Ca_{3-x}A_x)(Zr_{2-y}Fe_y)Fe_3O_{12}$ (A = Ce, Th, Gd) Ferrite Garnets as Revealed by Mössbauer Spectroscopy and the Rietveld Analysis

Academician V. S. Urusov*[1], V. S. Rusakov[1], Yu. K. Kabalov[1], and S. V. Yudintsev[2]

Received June 24, 2004

One of the most promising matrices for burying actinide-containing highly radioactive wastes is based on $A_3B_2T_3O_{12}$ phases with a garnet structure (the *Ia3d* space group). These have a three-dimensional framework formed by the TO_4 tetrahedra and BO_6 octahedra connected by common vertices. Large A cations in the dodecahedral oxygen environment (with a coordination number of 8) are embedded into this framework. Bivalent (Ca, Mn, Mg, Fe, Co, Cd) and trivalent (Y, rare-earth) cations may occupy the A positions. The octahedral B positions may be occupied by bivalent, trivalent (Fe, Al, Ga, Cr, Mn, In, Sc, Co, V), tetravalent (Zr, Ti, Sn, Ru), and even by pentavalent (Nb, Ta, Sb) cations. The centers of T tetrahedra most often contain tetravalent (Si, Ge, Sn) cations, but can also contain trivalent (Al, Ga, Fe) and pentavalent (N, P, V, As) cations. In addition, as a result of the isovalent and heterovalent isomorphism of both natural and synthesized phases with garnet structure, the formation of binary and multicomponent solid solutions occurs. This especially favors mixing of different components of highly radioactive wastes into the structure of these solid solutions.

In natural silicate garnets, the content of actinides, uranium, and thorium does not exceed a few fractions of a percent [1]. Therefore, in contrast to the other potential matrices for highly radioactive wastes, e.g., pyrochlore, studying them does not allow us to estimate the solubility of actinides in the garnet structure and to choose the most appropriate compositions. This information can be based only on experimental data. From a crystal chemical standpoint, the occupation of B and T positions by relatively large low-valence cations should favor the entrance of tetra- and trivalent actinide elements into the A position of the garnet structure. The largest trivalent cation capable of occupying both octahedral and tetrahedral positions is Fe^{3+}. Therefore, ferrites are the substances with the largest framework dimensions formed by FeO_6 octahedra and FeO_4 tetrahedra (i.e., they have the largest lattice constants) and, hence, with the largest size of AO_8 dodecahedral positions. Indeed, it turned out that, in ferrite garnets with the compositions

$$(Ca_{2.5}Th_{0.5})Zr_2Fe_3O_{12},\quad (Ca_2Th)(ZrFe)Fe_3O_{12}, \text{ and } (Ca_{2.5}Ce_{0.5})Zr_2Fe_3O_{12},$$

synthesized at temperatures T = 1050°C and 1200°C and at a pressure of 1 atm in ambient air [2], Th^{4+} and Ce^{4+} occupy up to one-third of the A-type positions. The ionic radii (according to Shannon and Prewitt [3] for a coordination number of 8) for such radioactive elements as Pu^{4+} (0.96 Å), Np^{4+} (0.98 Å), and U^{4+} (1.00 Å) lie within the range between the ionic radii of Th^{4+} (1.05 Å) and Ce^{4+} (0.97 Å). Therefore, we can expect that it is possible to obtain in the laboratory ferrite garnets with a high content of these elements. Indeed, ferrite garnets with higher contents of rare-earth elements have already been synthesized [4, 5]. This provides an opportunity to introduce into such a structure trivalent Am and Cm with close values of ionic radii.

For further study of the structure and properties of synthetic garnets, we prepared several polycrystalline samples of ferrites containing Ce, Gd, and Th. The initial components of the furnace charge (oxides and calcium carbonate) were ground in an agate mortar down to a particle size of about 20–30 μm and pressed into tablets at 25°C and 200 MPa. Afterwards, the tablets were sintered at 1300°C for five hours. Studies of the reaction kinetics in experiments with different durations demonstrated that this time is sufficient to attain the equilibrium state (the absence of phase-composi-

[1] *Moscow State University, Vorob'evy gory, Moscow, 119899 Russia*

* *e-mail: urusov@geol.msu.ru*

[2] *Institute of the Geology of Ore Deposits, Petrography, Mineralogy, and Geochemistry (IGEM), Russian Academy of Sciences, Staromonetnyĭ per. 35, Moscow, 109017 Russia*

1028-3358/04/4912-0715$26.00 © 2004 MAIK "Nauka/Interperiodica"

tion changes). Four samples with the following chemical formulas were prepared:

(**1C**) $Ca_{2.5}Ce_{0.5}Zr_2Fe_3O_{12}$,

(**2C**) $Ca_{1.5}GdCe_{0.5}ZrFeFe_3O_{12}$,

(**1T**) $Ca_{2.5}Th_{0.5}Zr_2Fe_3O_{12}$, and

(**2T**) $Ca_{1.5}GdTh_{0.5}ZrFeFe_3O_{12}$.

The samples were studied using a Philips PW3040/00 X'Pert MPD x-ray diffractometer (CuK_α-line radiation, 40 kV voltage, 20–30 mA electric current, 2θ angular range with θ = 3–65 angular degrees, a measurement step of 0.01–0.02 angular degrees, and a pulse accumulation time of 0.5–5 s at each point). The compositions of the phases were determined by a JCM-5300 + Link ISIS scanning electron microscope (25 keV, 1 nA, 1–3-μm probe diameter, 100-s pulse accumulation time). It was found that samples with similar stoichiometries have approximately the same phase compositions: the **1C** and **1T** samples consist of garnet and a small amount of calcium zirconate with a perovskite-type structure. The **2T** sample seems to contain a mixed ferrite of calcium and rare-earth elements, in addition to the prevailing garnet phase.

A more detailed study of the structure and phase composition of the described samples was performed on the basis of the Mössbauer effect and the Rietveld method. The latter is the full-profile analysis of the x-ray diffraction spectra carried out for powder specimens. An analysis of the valence and structural states of iron atoms was performed for ^{57}Fe nuclei in the "absorption" geometry at room temperature using an electrodynamic Mössbauer spectrometer in the constant-acceleration mode. We used ^{57}Fe nuclei in the Rh matrix with an activity equal to several mCi. The experimental spectra were studied by the methods of model analysis [6]. The reconstruction of distribution functions for the parameters characterizing the hyperfine structure of partial spectra [7] was performed on the basis of the SPECTR and DISTRI codes included into the MSTools program package [8].

The ^{57}Fe Mössbauer spectra nuclei in the samples under study were of the paramagnetic type and consisted of several quadrupole doublets with significantly different values of the quadrupole displacement ε for the spectral components and values of the Mössbauer line isomer shift δ. Initially, to solve the problem of the number of resonance lines and their distribution over partial components for all observed spectra, we reconstructed the distribution functions $p(v)$ for the positions of separate resonance lines. We assumed that the natural width Γ_τ of the Mössbauer line equals 0.097 mm/s and that, for the radiation source, it is 0.11 mm/s (according to its technical specifications). The analysis of the results concerning the reconstruction of the $p(v)$ functions allowed us to choose rather reasonable models and initial values of the varied parameters in order to refine the distribution functions $p(\varepsilon)$ for each of partial spectra and for further model analysis. The garnets under study represent phases of variable composition and relate to the so-called locally inhomogeneous systems [8]. Therefore, the model analysis of the observed spectra was performed under the assumption of the existence of two independent asymmetric quadrupole doublets with components of the same intensity and for sufficiently intense partial spectra. For the same reason, in the course of reconstructing the distribution function $p(\varepsilon)$ for quadrupole displacements in these spectra, we

Table 1. Results of the model analysis (SPECTR) and of an analysis of the reconstructed distribution functions for the parameters characterizing the hyperfine structure of the ^{57}Fe Mössbauer spectra (DISTRI) nuclei in the garnet samples under study (the standard deviation for statistical errors is indicated in brackets)

Sample	Position	SPECTR					DISTRI			
		I, %	δ, mm/s	ε, mm/s	Γ_1, mm/s	Γ_2, mm/s	*I*, %	δ, mm/s	ε, mm/s	Γ_1, mm/s
1T	T1	63.8(6)	0.248(2)	0.486(2)	0.288(3)	0.291(3)	60.0(1.5)	0.246(2)	0.474(15)	0.084(5)
	T2	35.1(6)	0.243(2)	0.648(2)	0.288(3)	0.313(4)	38.8(1.9)	0.245(3)	0.626(20)	0.109(21)
	T	1.1(1)	0.238(12)	0.062(12)	0.220(20)		1.6(3)	0.224(10)	0.071(12)	0.097(5)
2T	T	76.2(5)	0.192(1)	0.556(1)	0.349(2)	0.374(3)	78.1(9)	0.193(9)	0.551(26)	0.195(8)
	B	23.8(5)	0.404(3)	0.215(3)	0.303(5)		21.9(9)	0.396(3)	0.205(18)	0.169(10)
1C	T1	61.2(6)	0.181(2)	0.468(2)	0.291(2)	0.301(3)	62.6(1.5)	0.180(4)	0.463(18)	0.096(6)
	T2	30.5(5)	0.193(2)	0.686(2)	0.291(2)	0.296(4)	30.3(1.2)	0.195(6)	0.689(21)	0.086(8)
	B	8.3(3)	0.439(11)	0.286(11)	0.304(3)		7.1(6)	0.451(60)	0.262(32)	0.176(36)
2C	T	80.1(9)	0.201(2)	0.552(2)	0.382(3)	0.409(6)	81.5(9)	0.206(3)	0.531(29)	0.212(6)
	B	19.9(9)	0.422(6)	0.231(6)	0.285(6)	0.326(12)	18.5(9)	0.404(4)	0.217(27)	0.170(12)

Note: The isomer shift δ of the Mössbauer lines is taken with respect to the position of the α-Fe line.

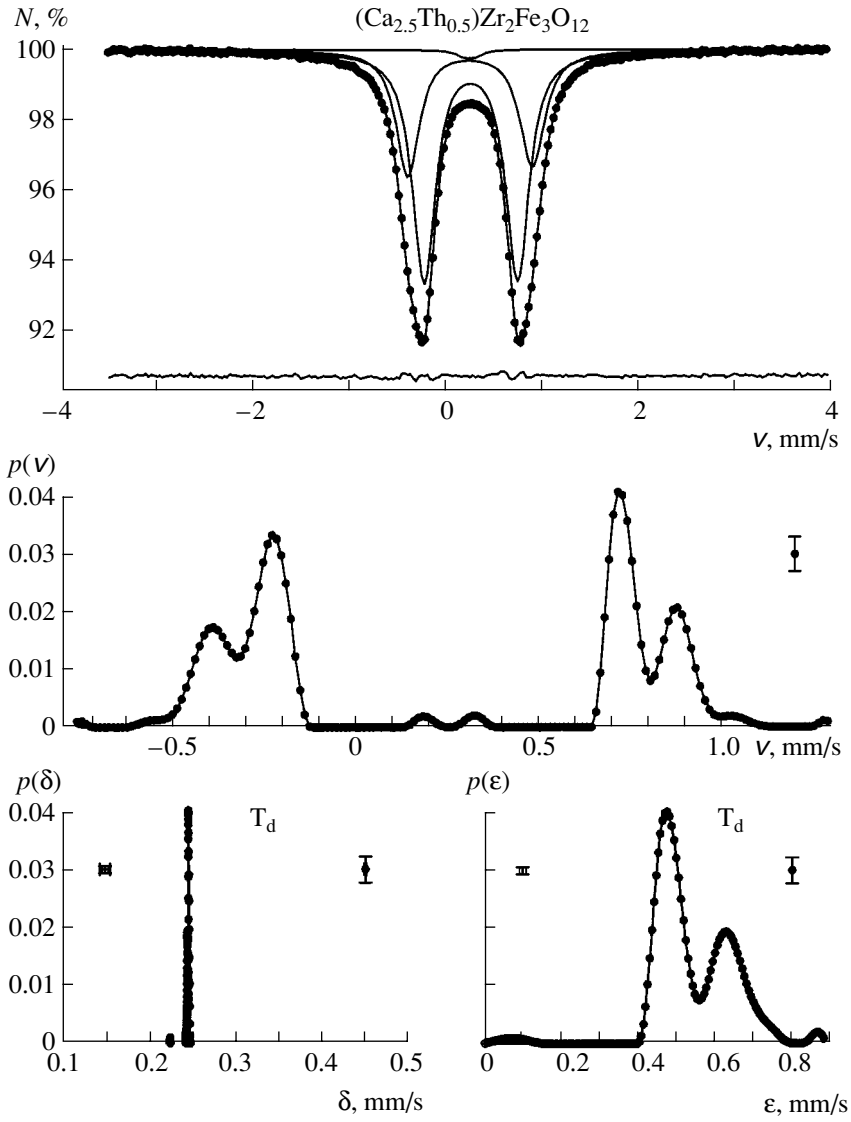


Fig. 1. Results of the model analysis and of the reconstruction of the distribution functions $p(V)$ for the positions of a separate resonance line and for the quadrupole displacement $p(\varepsilon)$ of the Mössbauer spectrum of ^{57}Fe nuclei in $(Ca_{2.5}Th_{0.5})Zr_2Fe_3O_{12}$ ferrite garnet (**1T** sample).

performed a search for a linear correlation between the quadrupole displacement ε and the isomer shift δ of the Mössbauer line [8].

In Figs. 1 and 2, as examples, we present for comparison the results of all the processing stages for the spectra of garnet samples (**1T**) $(Ca_{2.5}Th_{0.5})Zr_2Fe_3O_{12}$ and (**2T**) $(Ca_{1.5}GdTh_{0.5})(ZrFe)Fe_3O_{12}$. For the **1T** sample, the distribution function $p(V)$ definitely indicates the existence of three partial spectra in the form of quadrupole doublets embedded in each other with nearly the same isomer shifts δ of the Mössbauer line: two intense doublets and one much less intense doublet with the small quadrupole displacement ε. Similar results corresponding to two high-intensity doublets were also obtained for the Ce-containing **1C** sample ($(Ca_{2.5}Ce_{0.5})Zr_2Fe_3O_{12}$). For the **2T** sample containing Th and Gd, the distribution function $p(V)$ exhibited two clearly pronounced partial spectra with significantly different values of intensity I, isomer shift δ, and quadrupole displacement ε. In this case, the more intense quadrupole doublet is characterized by a smaller isomer shift and larger quadrupole displacement ε. The same situation was also observed for the **2C** sample containing Gd in addition to Ce:

$$(Ca_{1.5}GdCe_{0.5})(ZrFe)Fe_3O_{12}.$$

Using the reconstruction of the distribution functions for quadrupole displacements in the partial spectra and the subsequent model analysis, we managed to determine the values of the relative intensity I and of the parameters characterizing the hyperfine structure of the spectra under study. In Table 1, we present the results of

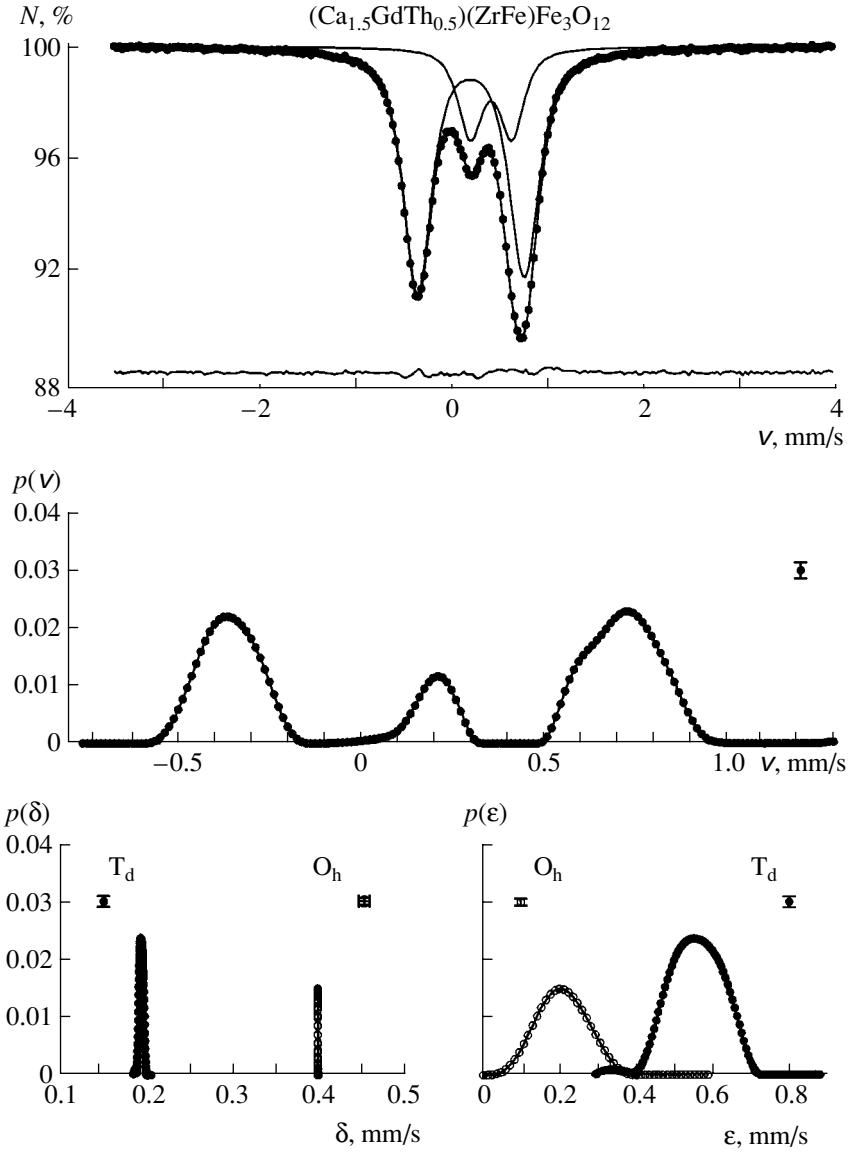


Fig. 2. Results of the model analysis and of the reconstruction of the distribution functions $p(V)$ for the positions of a separate resonance line and for the quadrupole displacement $p(\varepsilon)$ of the Mössbauer spectrum of ^{57}Fe nuclei in $(Ca_{1.5}GdTh_{0.5})(ZrFe)Fe_3O_{12}$ ferrite garnet (**2T** sample).

the model analysis for the spectra (isomer shift δ, quadrupole displacement ε, linewidths $\Gamma_{1,2}$ for the spectral components, and relative intensities I for each of partial spectra), as well as the characteristics of the reconstructed distribution functions (average values of the isomer shift $\bar{\delta}$ and of the quadrupole displacement $\bar{\varepsilon}$, widths $\Gamma_{p(\delta)}$ and $\Gamma_{p(\varepsilon)}$ of the distribution functions at half-height, and the value of the linear correlation for the parameters characterizing the hyperfine structure). Assuming that the values of the Debye temperature describing the vibration spectra of Mössbauer atoms at different crystallographic positions are close to each other, we have accepted that the ratio of the intensities for the corresponding partial spectra is equal to that of the occupation numbers for Mössbauer atoms at these positions.

The partial spectra with lower isomer shifts δ ($\delta \sim$ 0.18–0.25 mm/s) and with a larger quadrupole displacements ε ($\varepsilon \sim$ 0.47–0.69 mm/s) in each observed spectrum (Table 1) could be attributed to trivalent iron ions in the high-spin state at the positions with tetrahedral surrounding (T_d). Much larger isomer shift values ($\delta \sim$ 0.40–0.45 mm/s) and smaller quadrupole displacements ($\varepsilon \sim$ 0.21–0.29 mm/s) for other quadrupole doublets correspond to the high-spin state of Fe^{3+} ions at the B position with the octahedral (O_h) oxygen coordination.

For **1T** and **1C** samples, we reliably distinguish two different tetrahedral positions for iron ions. The ratio of

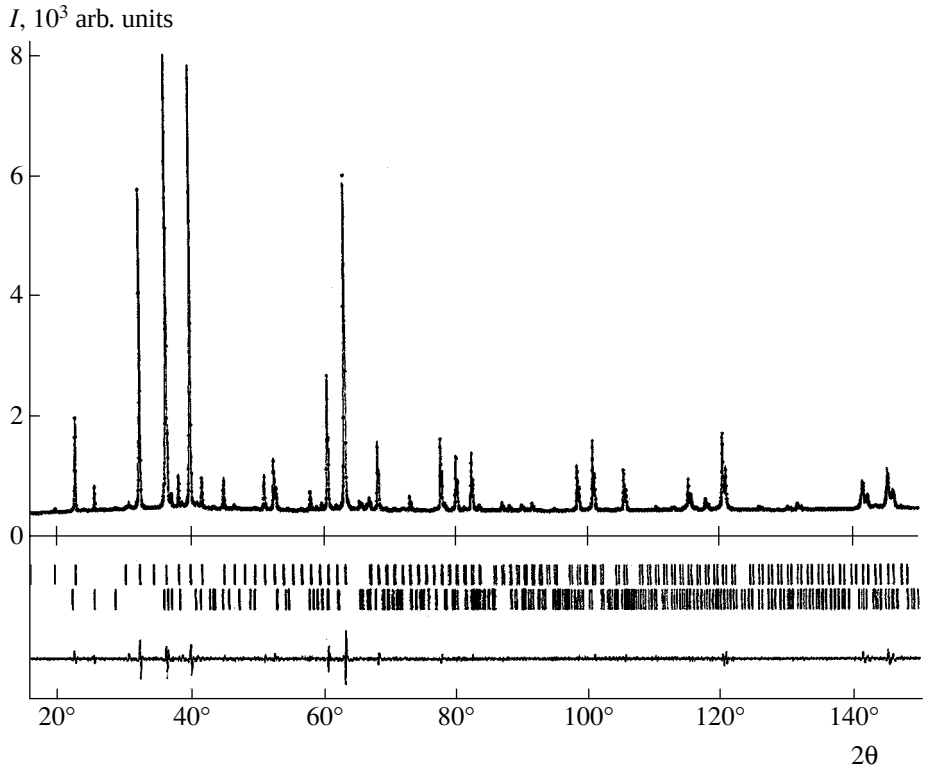


Fig. 3. Measured (solid line) and calculated (dots) x-ray diffraction spectra for the **1C** sample. The difference curve is shown in the bottom of the figure.

intensities for the corresponding Mössbauer partial spectra is close to two: 1.6–1.8 for the **1T** sample and 2.0–2.1 for the **1C** sample. The former does not contain iron at the octahedral position, whereas in the latter, the amount of iron is about 8%. The low-intensity quadrupole doublet ($I \simeq 1\%$) that appears in the spectrum of the **1T** sample, could be attributed to Fe^{3+} ions occupying tetrahedral positions of the phase, a small amount of which is contained in this sample.

In contrast, the **2T** and **2C** samples are characterized by a significant content of Fe^{3+} ions occupying octahedral positions. The relative intensities of the corresponding partial spectra are 22–24% for the **2T** sample and 19–20% for the **2C** sample. The linewidths for the components of the partial spectrum and widths $\Gamma_{p(\varepsilon)}$ for the distribution functions of the quadrupole displacement $p(\varepsilon)$ for the tetrahedral position significantly exceed those for the octahedral position (Table 1). This could be associated with the pronounced local inhomogeneity in the vicinity of Fe^{3+} ions occupying tetrahedral positions.

Thus, the data from Mössbauer spectroscopy, which were discussed above, present definite indications for different structural states characteristic of iron atoms in the two groups of samples under study: **1T**, **1C** and **2T**, **2C**. For the first pair of samples, we deal with two tetrahedral positions of Fe atoms with an occupation ratio of 2 : 1, whereas for the second pair, apart of a single type of tetrahedral positions there is an appreciable number of Fe atoms occupying octahedral positions. Note that the splitting of tetrahedral positions into two groups with an occupation ratio of 2:1 was also observed recently via Mössbauer spectroscopy in ferrite garnets of a similar composition with regard to Ce and Th but without additions of rare-earth elements [11]. The authors of [11] interpreted the unusual behavior of ferrite of such a composition as a result of an inhomogeneous local environment for two different kinds of tetrahedrally coordinated iron. However, it is clear that the random distribution of atoms substituting for each other at the A and B positions also results in randomness in the local environment around the T positions. However, this could cause only the line broadening of the corresponding quadrupole doublet, not the splitting of the spectra into two doublets. Such splitting implies a high degree of short-range order that it should inevitably stem from the long-range order in the structure under study. Therefore, the **1T** and **1C** samples from our sample set could have a crystal structure distinct from that characteristic of the **2T** and **2C** samples. These considerations encouraged us to turn once again to the x-ray diffraction technique.

After the repeated x-ray diffraction study of all four samples with an ADP-2 automated diffractometer (Co radiation, Fe filter), it was found that the **2C** and **2T** samples correspond to the cubic structure with the lattice constants $a = 12.622(5)$ Å and 12.660(5) Å, respec-

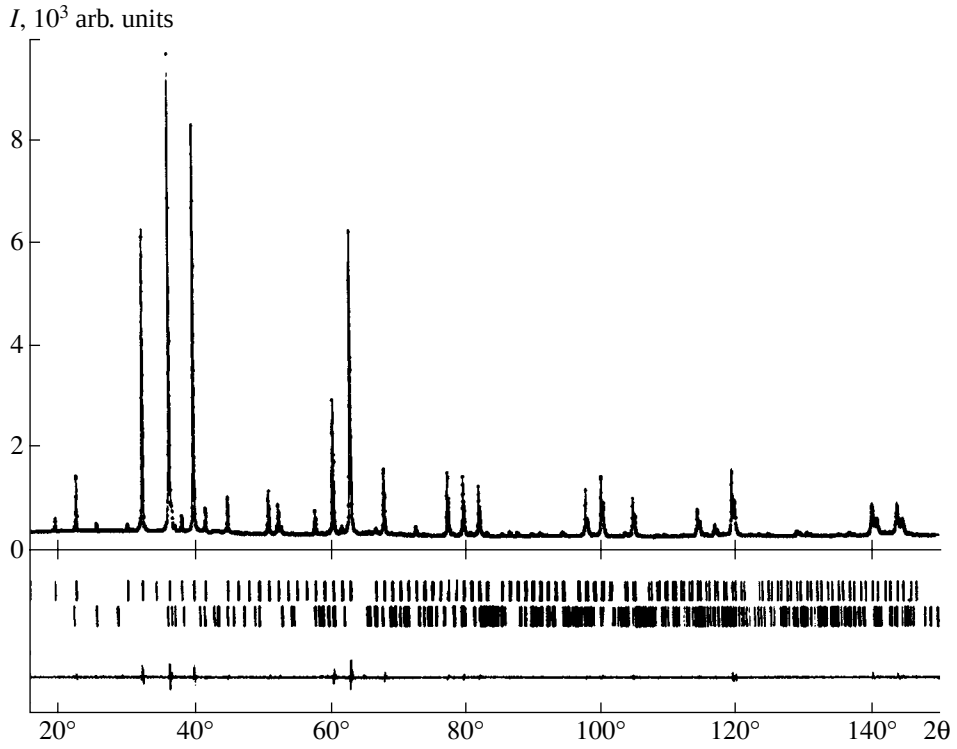


Fig. 4. Measured (solid line) and calculated (dots) x-ray diffraction spectra for the **1T** sample. The difference curve is shown in the bottom of the figure.

tively. On the basis of the quantitative x-ray phase analysis, the existence of small fractions with the perovskite structure having $CaZrO_3$ compositions of 8% in **1C** and 3% in **1T** samples was confirmed. Furthermore, the presence of some additional reflections allowed us to suggest a certain distortion of the cubic structure in **1C** and **1T** garnets. Therefore, the x-ray diffraction spectra of the **1C** and **1T** samples were recorded within the interval 2θ = 16.00–149.98 angular degrees with a scanning step equal to 0.02 degrees with a 20-s exposure at each point. All calculations concerning the refinement of the structural model were carried out using the 3.3 version of the WYRIET program package [12]. The approximation of the peak profile was performed on the basis of the Pearson-VII function. The refinement was made in the approximation of isotropic thermal vibrations of atoms and for several structural models: for cubic garnet (*Ia*3*d* space group) and for two of its tetragonal modifications with $I4_1/a$ and $I4/acd$ space groups. The initial atomic coordinates for the tetragonal modifications with $I4_1/a$ and $I4/acd$ space groups were taken from [13] and [14], respectively. After several refinement stages, it was found that the best agreement between the experimental and theoretical profiles is attained for the model of tetragonal garnet with the $I4/acd$ space group (Figs. 3 and 4). In Tables 2, 3, and 4, we present the basic results obtained in the course of such a refinement: the occupation numbers for different atomic positions, atomic coordinates, and the nearest interatomic distances for both the **1C** and **1T** samples.

Table 2. The results of the refinement by the Rietveld method for the crystal structures of the **1C** and **1T** samples

Parameter	**1C**	**1T**
a, Å	12.7130(3)	12.7546(3)
b, Å	12.7130(3)	12.7546(3)
c, Å	12.7153(5)	12.7588(3)
V, Å^3	2055.05(6)	2075.60(5)
Space group	$I4_1/acd$	$I4_1/acd$
2θ° range, angular degrees	16–149.98	16–149.98
Number of reflections $K(\alpha_1 + \alpha_2)$	684	684
Number of refined parameters	38	38
R_{wp}, %	5.14	5.34
R_F, %	2.96	2.10
*S**	1.17	1.07
*DWD***	1.50	1.84

* Goodness of fit.
** Statistics according to Durbin–Watson [15].

For comparison, we can indicate the main refinement criteria for other structural models describing the **1C** sample, which are as follows: cubic structure, *Ia*3*d*, $R_{wp} = 5.65\%$, $R_F = 3.14\%$, and $S = 1.29$; tetragonal

Table 3. Positions, temperature parameters, and interatomic distances in the structure of the **1T** sample

Atom	*x*	*y*	*z*	B_{iso}	Occupation number
Ca1	0.1215(8)	0	0.25	2.30(7)	1.52(2)
Th1	0.1215(8)	0	0.25	2.30(7)	0.48(2)
Ca2	0	0.25	0.125	0.7(1)	0.87(2)
Th2	0	0.25	0.125	0.7(1)	0.13(2)
Zr	0	0	0	1.02(3)	1.98(2)
Fe1	0.380(1)	0	0.25	2.5(1)	1.95(2)
Fe2	0	0.25	0.375	0.9(2)	1.00(2)
O1	0.2886(8)	0.0975(8)	0.1997(8)	0.2(3)	4.00
O2	0.1008(8)	0.193(1)	0.276(1)	0.5(3)	4.00
O3	0.191(1)	0.2885(8)	0.092(1)	0.7(3)	4.00

Interatomic distances, Å

Ca1–O1 = 2.53(3) × 2	Ca2–O2 = 2.40 × 4	Zr–O1 = 2.15(3) × 2
Ca1–O1' = 2.54(3) × 2	Ca2–O3 = 2.58 × 4	Zr–O2 = 2.10(3) × 2
Ca1–O2 = 2.53(4) × 2	Average Ca–O = 2.49	Zr–O3 = 2.14(4) × 2
Ca1–O3 = 2.46(4) × 2		Average Zr–O = 2.13
Average Ca–O = 2.48		
Fe1–O1 = 1.84(4) × 2	Fe2–O2 = 1.90(4) × 2	
Fe1–O3 = 1.75(5) × 2	Fe2–O2' = 1.90(4) × 2	
Average Fe–O = 1.82	Average Fe–O = 1.90	

Table 4. Positions, temperature parameters, and interatomic distances in the structure of the **1C** sample

Atom	*x*	*y*	*z*	B_{iso}	Occupation number
Ca1	0.124(4)	0	0.25	1.6(1)	1.48(2)
Ce1	0.124(4)	0	0.25	1.6(1)	0.52(2)
Ca2	0	0.025	0.125	0.6(2)	0.88(2)
Ce2	0	0.025	0.125	0.6(2)	0.12(2)
Zr	0	0	0	0.6(3)	2.00(2)
Fe1	0.378(5)	0	0.25	2.1(2)	1.92(2)
Fe2	0	0.25	0.375	0.6(2)	1.00(2)
O1	0.279(6)	0.094(5)	0.192(5)	0.4(4)	4.00
O2	0.098(4)	0.194(6)	0.276(5)	0.4(4)	4.00
O3	0.197(5)	0.288(4)	0.095(5)	0.6(5)	4.00

Interatomic distances, Å

Ca1–O1 = 2.43(5) × 2	Ca2–O2 = 2.40(4) × 4	Zr–O1 = 2.15(4) × 2
Ca1–O1' = 2.50(4) × 2	Ca2–O3 = 2.58(4) × 4	Zr–O2 = 2.09(3) × 2
Ca1–O2 = 2.52(5) × 2	Average Ca–O = 2.49	Zr–O3 = 2.14(4) × 2
Ca1–O3 = 2.47(4) × 2		Average Zr–O = 2.13
Average Ca–O = 2.48		
Fe1–O1 = 1.88(5) × 2	Fe2–O2 = 1.90(4) × 2	
Fe1–O3 = 1.75(5) × 2	Fe2–O2' = 1.90(4) × 2	
Average Fe–O = 1.82	Average Fe–O = 1.90	

structure, $I4_1/a$, $R_{wp} = 5.29\%$, $R_F = 4.38\%$, and $S = 1.21$. Comparing the refinement results listed in Table 2 for this sample, we can conclude that, according to all three accuracy criteria, the tetragonal structure with the *I*4/*acd* space group is significantly more advantageous than the other structures.

The idealized crystal-chemistry formulas that are the result of the refinement of the occupation numbers (taking into account possible errors) for different crystallographic positions, namely, the two positions A1 and A2, the single position B, and the two positions T1 and T2, are the following.

1C sample: $(Ca_{1.50}Ce_{0.50})^{A1}(Ca_{0.90}Ce_{0.10})^{A2}Zr^{B}_{2.00}(Fe_{1.93})^{T1}(Fe_{1.00})^{T2}O_{12}$,

1T sample: $(Ca_{1.50}Th_{0.50})^{A1}(Ca_{0.90}Th_{0.10})^{A2}Zr^{B}_{2.00}(Fe_{1.93})_{T1}(Fe_{1.00})^{T2}O_{12}$.

We can see that the gross compositions of both samples slightly differ from the initial ones that were determined on the basis of the recalculated results of chemical analyses. Indeed, there is a tendency toward a certain increase in the content of Ce and Th at the expense of the atomic percentage of Ca and Fe (at the T1 tetrahedral position with double occupancy). These elements may be absorbed by "parasitic phases," namely, perovskites and oxides existing in the samples. In addition, we should emphasize that the tetravalent elements obviously prefer position A1 over A2 and that there is also a possible small deficiency (of about 3%) of iron at the T1 position. In general, the atomic fraction of Ce and Th at the A positions is as high as 30%. It probably becomes slightly higher than in the initial furnace charge due to the increase in the number of A-type positions in the tetragonal modification compared to the perfect cubic structure of garnets. The absence of a tetragonal distortion in **2C** and **2T** samples containing the third element at the A position, namely, trivalent Gd, can be attributed to its intermediate character, which facilitates the mixing of bivalent and tetravalent elements at the same position, and to an increase in the configurational entropy (the entropy-driven stabilization of the multicomponent mixture).

ACKNOWLEDGMENTS

The work was supported by the Council of the President of the Russian Federation for Support of Young Russian Scientists and Leading Scientific Schools (project no. NSh-1955.2003.5) and by the Russian Foundation for Basic Research (project no. 02-05-64845).

REFERENCES

1. H. Jaffe, Am. Mineral. **36**, 133 (1951).
2. J. Ito and C. Frondel, Am. Mineral. **52**, 773 (1967).
3. R. D. Shannon and C. T. Prewitt, Acta Crystallogr. B **26**, 1046 (1970).
4. A. L. Gentile and R. Roy, Am. Mineral. **45**, 701 (1960).
5. S. Geller, Z. Kristallogr. **125**, 1 (1967).
6. V. I. Nikolaev and V. S. Rusakov, *Mössbauer Studies of Ferrites* (Moscow State University, Moscow, 1985).
7. V. S. Rusakov, Izv. Ross. Akad. Nauk, Ser. Fiz. **63**, 1389 (1999).
8. V. S. Rusakov, *Mössbauer Spectroscopy of Locally Inhomogeneous Systems* (Institute of Nuclear Physics, Kazakhstan National Nuclear Center, Almaty, 2000).
9. F. Menil, J. Phys. Chem. Solids **46**, 763 (1985).
10. R. V. Parish, in *Mössbauer Spectroscopy,* Ed. by D. P. E. Dickson and F. J. Berry (Cambridge Univ. Press, Cambridge, 1986).
11. B. G. Shabalin, E. V. Pol'shin, Yu. O. Titov, *et al.*, Miner. Zh. **25**, 41 (2003).
12. J. Schneider, in *Proceedings of IV Cryst. Int. Workshop on the Rietveld Method, Petten, 1989*, p. 71.
13. A. Nakatsuka, A. Yoshiasa, T. Yamanaka, *et al.*, Am. Mineral. **84**, 199 (1999).
14. D. T. Griffen, D. M. Hatch, W. R. Phillips, *et al.*, Am. Mineral. **77**, 399 (1992).
15. R. Hill and H. J. Flack, Appl. Crystallogr. **20**, 356 (1987).

Translated by K. Kugel

Doklady Physics, Vol. 49, No. 12, 2004, pp. 723–726. Translated from Doklady Akademii Nauk, Vol. 399, No. 6, 2004, pp. 753–756.

PHYSICS

On Irreversibility in Quantum Mechanics

N. V. Lunin* and V. I. Kogan**

Presented by Academician S.T. Belyaev June 22, 2004

Received June 22, 2004

In connection with certain physical problems, Ginzburg [1] noted that the physical causes of a monotonic increase in the entropy of closed systems, as well as a corresponding irreversibility, are not yet clearly understood. A usual approach to irreversibility in quantum mechanics is as follows. Dynamic equations, including the Schrödinger equation, are time reversible. Based on this circumstance, one concludes that postulated quantum-mechanical schemes cannot lead to irreversibility in closed systems. For this reason, either a transition to open systems or a substantial change in the mathematical foundations of quantum mechanics and transition to the equipped Hilbert space is proposed in order to include irreversibility in quantum mechanics. Other proposals involve the inclusion of nonlinear and complex terms in the Schrödinger equation.

Space–time symmetries, which are mathematically expressed in terms of the theory of groups, i.e., sets with one operation where each element has an inverse element, are responsible for exact conservation laws. The Noether theorems establish a one-to-one correspondence [2] between conservation laws and group-theoretical requirements imposed on physical theory, and, owing to the reversibility of equations, all propagators (matrices transforming solutions from one space–time point to another) have inverse operators. Therefore, the satisfaction of conservation laws that follow from reversible equations leads to group-theoretical requirements for propagators, in particular, to the existence of inverse propagators; i.e., the reversibility of equations and the group-theoretical construction of a physical theory are mutually dependent [3].

The description of a physical system that includes irreversible equations is a sufficient condition of its irreversible evolution. Moreover, its description by only reversible equations is obviously a necessary condition of its reversible evolution. Is the latter condition also a sufficient condition of reversible evolution?

The noncommutativity of operators (propagators) transforming solutions is one of the remarkable features of quantum mechanics. This feature is naturally manifested in the evolution of observables. For this reason, it is of interest to determine whether irreversibility is connected with noncommutativity and, since the latter is also involved in reversible equations, whether it can lead to the irreversibility of evolution.

This work is aimed at determining the connection of irreversibility with noncommutativity and the conditions for the reversible and irreversible evolution of a closed physical system described only by reversible equations. These problems are analyzed on the basis of the existence of a commutative binary composition of elements on noncommutative groups and the corresponding possibility of a group-theoretical scheme of quantum mechanics [4, 5].

Let the system be described by the Schrödinger equation with a real potential depending on coordinates and time. Expressing the wave function and its gradient in terms of the spinor Φ = column $\|\Phi_+, \Phi_-\|$ [5], we pass from the Schrödinger equation to two equations for $\Phi_\pm$, i.e., to the spinor representation of the Schrödinger equation

$$\sigma_0 \nabla \Phi + \frac{\mathbf{u}}{2k}(\sigma_3 + i\sigma_2)\Phi' = \left\{ ik\mathbf{u}\sigma_3 + \frac{\nabla k}{2k}\sigma_1 + \frac{\mathbf{u}}{2k}\frac{k'}{2k}(\sigma_3 + i\sigma_2) \right\}\Phi. \quad (1)$$

Here, the prime stands for the time derivative; $k^2(\mathbf{r}, t) = \frac{2m}{\hbar^2}\{E_0 - U(\mathbf{r}, t)\}$; σ_0 is the identity matrix; σ_i, $i = 1, 2, 3$, are the Pauli matrices; E_0 is the energy at the point $(\mathbf{r}_0, t_0)$; and $2m = \hbar = 1$. We use the Feynman scheme of quantum mechanics with the stipulation that the transition amplitude along each path is a multiplicative integral [6], i.e., the limit of the product of matrix transformations, including noncommutativity, and the total amplitude is their composition over all paths that conserves the group properties [4, 5]. In this scheme, we treat $\mathbf{u}(\mathbf{r})$ in Eq. (1) as a set of unit vectors that specify

* *Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603600 Russia*
e-mail: n.v.lunin@rambler.ru

** *Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia*

1028-3358/04/4912-0723$26.00 © 2004 MAIK "Nauka/Interperiodica"

paths from the initial to the final point and that satisfy the condition $\nabla \mathbf{u} = 0$.

Expressing the solution as $\Phi(\mathbf{r}, t) = Q\,\Phi(\mathbf{r}_0, t_0)$, we obtain the multiplicative derivative [6] of the matrix Q in the form

$$\left\{\nabla Q + \frac{\mathbf{u}}{2k}(\sigma_3 + i\sigma_2)Q'\right\}Q^{-1} = \mathbf{P}, \qquad (2)$$

where $\mathbf{P}$ is the matrix appearing in braces in Eq. (1). Its form provides $Q = \exp\{\mathbf{q}\boldsymbol{\sigma}\}$ [5], where $\mathbf{q}$ is a complex vector and, therefore, $Q \in SL(2, C)$.

A complete system of observables for the physical system described by the time-dependent Schrödinger equation has the same form as for the time-independent case [5]. This system consists of two scalars and two vectors, where the latter quantities are linear in the gradient of the wave function. Their amplitudes are four Hermitian forms $j_s = \Phi^+\sigma_s\Phi$, $s = 0, 1, 2, 3$, with the completeness condition $j_0^2 = j_1^2 + j_2^2 + j_3^2$. This condition is equivalent to the condition that the Hermitian density matrix for pure states [7] $J = \Phi^+ \otimes \Phi \equiv \frac{1}{2}\sum_{s=0}^{3}\sigma_j j_s$ has a zero determinant. For the current

$$\mathbf{j}_3 = i(\psi\nabla\psi^* - \psi^*\nabla\psi) = (\Phi^+\sigma_3\Phi)\mathbf{u},$$

where ψ is the solution of the Schrödinger equation, Eq. (1) and the Hermitian conjugate equation lead to the usual continuity equation

$$\nabla\mathbf{j}_3 + \frac{\partial\rho}{\partial t} = 0,$$

where $\rho = \psi\psi^* = \frac{j_0 + j_1}{2k}$. This equation, obtained from Eq. (1), is valid for all propagators $Q \in SL(2, C)$ that lead to the satisfaction of conservation laws. A partial propagator, i.e., a propagator for one path, belongs to the group and thereby has the inverse operator for any path, which agrees with the reversibility of the Schrödinger equation.

Since they are experimental facts, both irreversibility and reversibility must be defined only in terms of observables. For this reason, following [8], we accept that evolution is reversible if all observables in the time-reversed process pass through the same sequence as in the direct process but in the reverse order. Otherwise, evolution is irreversible. Owing to the reversibility of equations, conservation laws valid for the direct process must obviously be satisfied in the reverse process. Thus, all propagators for the reverse and direct processes must belong to the same group.

When the system evolves from t_0 to t, the density matrices for pure states form a sequence $J(t_0), \ldots, J(t)$, and any density matrix has no inverse matrix. Let spinors $\Phi(t_1)$ and $\Phi(t_2)$, along with the corresponding density matrices for pure states $J(t_1)$ and $J(t_2)$, exist at the respective arbitrary times t_1 and t_2 from the sequence $t_0 \ldots t_1 \ldots t_2 \ldots t$. These spinors are related by the matrix $Q(t_2, t_1)$ that is determined by Eq. (2) and that belongs to the $SL(2, C)$ group. Density matrices for pure states are written in the form [7]

$$J(t_1) = \frac{1}{2}\sum_{s=0}^{3}\sigma_s\{\Phi^+(t_1)\sigma_s\Phi(t_1)\}, \qquad (3)$$

$$J(t_2) = \frac{1}{2}\sum_{s=0}^{3}\sigma_s\{\Phi^+(t_1)Q^+(t_2, t_1)\sigma_s Q(t_2, t_1)\Phi(t_1)\}. \qquad (4)$$

Let the evolution of the system be described by only one sequence of (noncommutative) propagators $Q(t_0, t_0) \ldots Q(t_1, t_0) \ldots Q(t_2, t_0) \ldots Q(t, t_0)$ and a corresponding sequence of density matrices for the pure states $J(t_0) \ldots J(t_1) \ldots J(t_2) \ldots J(t)$. Under time reversal, a spinor is transformed by inverse matrices due to the reversibility of the Schrödinger equation, and $Q(t_1, t_2) = Q^{-1}(t_2, t_1)$. Since transformations of the spinor are multiplicative, the spinor is transformed in the direct and following reverse process $t_1 \longrightarrow t_2 \longrightarrow t_1$ as

$$\Phi(t_1) \to Q(t_2, t_1)\Phi(t_1) \to Q(t_1, t_2)Q(t_2, t_1)\Phi(t_1)$$
$$= Q^{-1}(t_2, t_1)Q(t_2, t_1)\Phi(t_1) = \Phi(t_1),$$

and the density matrix for pure states is transformed as $J(t_1) \longrightarrow J(t_2) \longrightarrow J(t_1)$; i.e., all observables at the end of the process take the values that they had at the beginning. Thus, if a process is completely described by one sequence of propagators, it is reversible.

Let the evolution of the system be described by more than one sequence of partial operators. It is sufficient to consider two sequences denoted as A and B in the $t_1 \longrightarrow t_2$ process. The total propagator, which includes both paths and provides for the construction of the density matrix, is a composition of partial propagators that satisfies the same conservation laws as A and B separately. Therefore, this propagator belongs to the same group that contains A and B. Since both evolution paths are equivalent, the total propagator must have not only the group properties attributed to conservation laws but also properties associated with discrete symmetries, particularly under permutations.

The reversibility of the Schrödinger equation means that its solution is equivalent to the solution of the complex conjugate equation and that the construction of the propagator for the spinor form of the Schrödinger equation is equivalent to the construction of the propagator for the corresponding spinor equation for the complex conjugate Schrödinger equation. Time reversal in the Schrödinger equation (along with the transition to complex conjugate solutions) corresponds to the change $t \longrightarrow -t$. In this case, the propagator Q is transformed to Q^{-1} for the reverse process. Therefore, the propaga-

tors Q and Q^{-1} (both inevitably exist on the group) for the direct and reverse process are equivalent in a manner analogous to that in which ψ and ψ^* are equivalent under time reversal. This means that the total propagator either remains unchanged or changes to the inverse propagator.

The following four compositions of propagators on noncommutative multiplicative groups satisfy these requirements [4, 5, 7, 9]:

$$M = \{(AB^{-1})^{1/2}B\}^2, \quad D = \{(AB)^{1/2}B^{-1}\}^2, \qquad (5)$$

$$T = (AB^{-2}A)^{1/2}A^{-1}B, \quad T' = (AB^2A)^{1/2}A^{-1}B^{-1}. \qquad (6)$$

If $A \longleftrightarrow B$, then $M \longrightarrow M$, $D \longrightarrow D^{-1}$, $T \longrightarrow T^{-1}$, and $T' \longrightarrow (T')^{-1}$ (the proof is simple and, thus, omitted). Since products and real powers of the elements of the group also belong to the group, all these compositions belong to the same group that contains A and B.[1]

Expressions (5) and (6) have a geometric interpretation in spaces with nonzero Gaussian curvature [4, 5]. Expressing the matrices in Eqs. (5) and (6) in the exponential form $A = \exp(\mathbf{a}\boldsymbol{\sigma})$, etc., and mapping the vectors **a**, **b**, **m**, … as geodesic vectors, one can show that the vectors **m** and **d** are the symmetric and antisymmetric diagonals, respectively, of the parallelogram drawn on the adjacent vectors **a** and **b**, and that each of the vectors **t** and **t**' is orthogonal to the vectors **a** and **b** and corresponds to the area of adjacent parallelograms.

If the vectors **a** and **b** are small, it follows from the expansion of the exponential representation of matrices in Eqs. (5) and (6) that

$$\mathbf{m} \simeq \mathbf{a} + \mathbf{b}, \quad \mathbf{d} \simeq \mathbf{a} - \mathbf{b}, \quad \mathbf{t} \simeq i[\mathbf{ba}] \simeq -t'.$$

The above analysis shows that the expressions for M, D, T, and T' express the non-Euclidean superposition principle, which includes the noncommutativity of propagators, leads to the satisfaction of conservation laws due to group properties, and changes to the Euclidean superposition principle in the small region of the space of the logarithms of propagators.

Under time reversal, all partial propagators change to the respective inverse propagators due to the reversibility of the Schrödinger equation. It is easy to show that the compositions M and D also change to the inverse compositions when $A \longrightarrow A^{-1}$ and $B \longrightarrow B^{-1}$. Therefore, the transformation of the spinor in the process $t_1 \longrightarrow t_2 \longrightarrow t_1$ by means of these propagators leads to the initial spinor.

For this reason, the compositions M and D describe only the reversible component of the evolution of the system.

Let us prove that the composition changes as $T \longrightarrow T$; i.e., $(A^{-1}B^2A^{-1})^{1/2}AB^{-1} = (AB^{-2}A)^{1/2}A^{-1}B$ when $A \longrightarrow A^{-1}$ and $B \longrightarrow B^{-1}$. Multiplying this equality by B and A^{-1} from the right and taking the square of the resulting equality, we obtain

$$A^{-1}B^2A^{-1}$$
$$= (AB^{-2}A)^{1/2}A^{-1}BA^{-1}(AB^{-2}A)^{1/2}A^{-1}(BBA^{-1})$$
$$= (AB^{-2}A)^{1/2}(AB^{-2}A)^{-1}(AB^{-2}A)^{1/2}(A^{-1}B^2A^{-1})$$
$$= A^{-1}B^2A^{-1}.$$

The composition T' has the same property.

Both discrete symmetries are conserved for small **a** and **b**. Compositions (6) can be treated as commutators on the multiplicative group that differ from the group theoretical expression $B^{-1}A^{-1}BA$. In this case, if the matrices A and B are commutative, T and T' are equal to the unit matrix.

Thus, there are compositions of alternative propagators that ensure the satisfaction of conservation laws and that do not change under time reversal and the corresponding inversion of partial propagators; i.e., they are transformed as $1 \longrightarrow T \longrightarrow T^2$ in the process $t_1 \longrightarrow t_2 \longrightarrow t_1$. At the end of the process $t_1 \longrightarrow t_2 \longrightarrow t_1$, the density matrix for pure states that appears in Eq. (4) and that is constructed by means of one such propagator takes the form

$$\tilde{J}(t_1) = \frac{1}{2}\sum_{s=0}^{3}\sigma_s\{\Phi^+(t_1)(T^2)^+\sigma_sT^2\Phi(t_1)\}. \qquad (7)$$

In the general case, it does not coincide with the initial form. Therefore, the process is irreversible.

Nevertheless, there are conditions under which such total propagators lead to reversibility. A comparison of Eqs. (3) and (7) shows that this is the case when

$$(T^2)^+\sigma_sT^2 = \sigma_s, \quad s = 0, 1, 2, 3. \qquad (8)$$

The violation of this condition can be treated as a sufficient condition of irreversibility. The satisfaction of this condition can be treated as a sufficient condition of reversibility, which is the case when A and B are commutative.

In systems described by a large number of partial propagators, compositions of the type T for which condition (8) is satisfied can coexist with compositions for which condition (8) is violated. This means that reversible processes can coexist with irreversible processes in systems described by only reversible equations.

In connection with the above approach to irreversibility, we present two conclusions from [10]. The first conclusion is that irreversibility expressed by the time arrow is a statistical property. It cannot be introduced in terms of individual paths or wave functions. Therefore, it demands a radical withdrawal from Newtonian mechanics or from orthodox quantum mechanics based on concepts of the individual path or wave function.

[1] The permutation properties of the expressions in Eqs. (5) and (6) are valid for noncommutative nonsingular matrices A and B of any order, as well as for any real power of these compositions.

The second conclusion is that the main assumption that we have introduced here is the statement that the space with zero Gaussian curvature, similar to Minkowski space, does not contain entropy. Both conclusions are explained in this approach: irreversibility is absent either if only one path exists or if the space of the logarithms of the propagators has zero Gaussian curvature; i.e., all propagators are commutative.

The noncommutativity of the propagators A and B is manifested in the acollinearity of the geodesic vectors **a** and **b** in non-Euclidean space, and the vector **t** in the composition T corresponds to the area of a closed contour, which is a parallelogram in the case under consideration. Expression (7) makes it possible to treat the vector **t** as an irreversibility measure. Since the area mentioned above is related to the Berry phase, the latter can serve as an irreversibility measure.

To determine the physical causes of irreversibility, we analyze the expression for the matrix **P** in Eq. (2) that determines the forms of the propagators. If $\nabla k = k' = 0$, **P** and the corresponding propagators are diagonal and commutative. Therefore, T is equal to the identity matrix, and irreversibility is absent. If $\nabla k \neq 0$ and/or $k' \neq 0$, A and B can be noncommutative, and the process can be irreversible. For this reason interaction, in particular, collisions, should be treated as a necessary, although insufficient, physical cause of irreversibility. For example, if interaction is present but $A = B$, the matrix T is a unit matrix and irreversibility is absent.

As a physical system where reversible and irreversible processes occur simultaneously, we consider a gedankenexperiment with two infinitely small (pointlike) apertures at the interface between two media. In the framework of one (reversible) composition M, this example was considered in [5], where the satisfaction of conservation laws was analyzed. The coexistence of reversible and irreversible processes in such a system is not surprising, because propagation under these conditions is described on the pseudounitary group $SU(1, 1)$, and one of the necessary conditions of reversibility is the unitarity of transformations, which follows from Eq. (8). Due to symmetry, the partial propagators A and B are obviously identical for all points on the symmetric z axis (see Fig. 1 in [5]). Therefore, $T = 1$. We note that this condition is also satisfied for all points on the hyperboloids of revolution that correspond to the maxima of the interference pattern. Therefore, propagation from a source to the indicated points is reversible. The propagators A and B for all other points, as follows from their explicit form in [5], are noncommutative in the presence of a potential jump. Therefore, condition (8) is not satisfied for them, and this process is thereby irreversible for them.

A joint consideration of the physical causes of irreversibility, i.e., interactions, as well as its geometric causes show that the noncommutativity of propagators leads to the existence of an observable that can serve as an irreversibility measure in the evolution of the physical system. This is the only nonzero positively defined eigenvalue of the density matrix for pure states and is a functional that increases with time, is bounded, and that rapidly reaches the limiting value with the enhancement of interactions.

Since the discrete symmetries of the compositions M, D, T, and T' are valid for nonsingular matrices A and B of any order, the above results are not restricted to only the Schrödinger equation—they can be generalized for more complex systems.

ACKNOWLEDGMENTS

We are grateful to S.T. Belyaev, Yu.L. Klimontovich, G.A. Kotel'nikov, V.S. Lisitsa, V.I. Man'ko, and L.A. Shelepin for their attention to the work, discussions, and comments.

REFERENCES

1. V. L. Ginzburg, Usp. Fiz. Nauk **169** (4), 419 (1999).
2. P. J. Olver, *Applications of Lie Groups to Differential Equations* (Springer, New York, 1986; Mir, Moscow, 1989).
3. I. A. Malkin and V. I. Man'ko, *Dynamical Symmetries and Coherent States of Quantum Systems* (Nauka, Moscow, 1979) [in Russian].
4. N. V. Lunin, Preprint No. 354, IPF RAN (Inst. of Applied Physics, Russian Academy of Sciences, 1994).
5. N. V. Lunin, Dokl. Akad. Nauk **368** (3), 323 (1999) [Dokl. Phys. **44**, 595 (1999)].
6. F. R. Gantmakher, *The Theory of Matrices* (Nauka, Moscow, 1988; Chelsea, New York, 1959).
7. N. V. Lunin, in *Proceedings of the XXII International Workshop on HEP and QFT, Protvino, Russia, 1999*, p. 272.
8. M. A. Leontovich, *Introduction to Thermodynamics* (Nauka, Moscow, 1983) [in Russian].
9. N. V. Lunin, Vestn. Nizhegor. Univ. Mat. Model. Opt. Upravl., No. 1 (25), 120 (2002).
10. I. Prigogine and I. Stengers, *Time, Chaos, Quantum* (Freeman, New York, 1994; Progress, Moscow, 1999).

Translated by R. Tyapaev

Doklady Physics, Vol. 49, No. 12, 2004, pp. 727–729. Translated from Doklady Akademii Nauk, Vol. 399, No. 6, 2004, pp. 757–759.

PHYSICS

The Ecton Mechanism of Breakdown of Solid Dielectrics

Academician G. A. Mesyats

Received July 26, 2004

In spite of a long history of investigations [1–3], the mechanism of breakdown of solid dielectrics remains unclear up to now. One of the reasons for studies of this phenomenon is its wide application in modern technology [4]. The breakdown of solid dielectrics is affected by many factors such as characteristics of the voltage applied, properties of the dielectric, the medium in which the breakdown occurs, etc. These factors also include conditions under which the electrode and the dielectric make contact. The surface of a metal electrode is usually irregular [5]. On this surface, there exist microprotrusions at which the electric field multiply enhanced. As for cathode, by virtue of this enhancement, the electric current density j of the field emission from the microprotrusions is multiply increased. This results in the electrical explosion of the microprotrusions, which is accompanied by explosion electron emission [5]. The electron emission of this type occurs in the form of separate short-time (10^{-9}–10^{-8} s) bursts each containing 10^{11}–10^{12} electrons. The author, who pioneered in the discovery of this phenomenon, called these bursts ectons [6, 7]. The volume of the cathode region in which this microexplosion occurs is close to 10^{-12} cm^3, and the energy density prior to the microexplosion is about 10^4 J g^{-1}. As a result of the microexplosion, a plasma jet is formed that moves from the cathode surface at a velocity on the order of 10^6 cm s^{-1}. It has been found that in vacuum discharge, without doubt, all these processes take place [5].

In [8], it has been established that the microexplosions at the metal surface can be initiated not only by virtue of the field-emission current but also by the displacement current. The displacement current is caused by the surface flashover at the metal–solid-dielectric–medium triple junction formed by the contact of a microprotrusion with the solid dielectric. In this case, the medium can be vacuum, gas, or liquid dielectric.

We consider the breakdown of a solid-dielectric plate of thickness d situated between a pin electrode and a plane electrode (Fig. 1). We assume that on the pin surface, there exists a cylindrical microprotrusion of radius r, which holds the pin in contact with the surface of the solid dielectric.

Let a pulse of high voltage U of amplitude U_a with a perfectly steep leading edge (which corresponds to the voltage drop U_a) be applied to the pin electrode. Let the amplitude U_a be high enough to initiate the discharge over the dielectric surface.

At the initial stage ($t < 10^{-7}$ s), this discharge is characterized by a continuous luminescence of the discharge plasma [5, 9]. Later, on this surface, separate local plasma filaments (Lichtenberg figures) are formed [9]. We now analyze just this initial stage of the process.

Radius R and area S of the luminous plasma zone is determined by the formulas

$$R = vt, \quad S = \pi v^2 t^2, \tag{1}$$

where v is the velocity of the plasma motion over the dielectric surface. If $R > d$, then the dynamic capacity $C(t)$ resulting from the plasma motion is

$$C = \frac{\pi\varepsilon_0\varepsilon v^2 t^2}{d}. \tag{2}$$

The generated displacement current is determined by

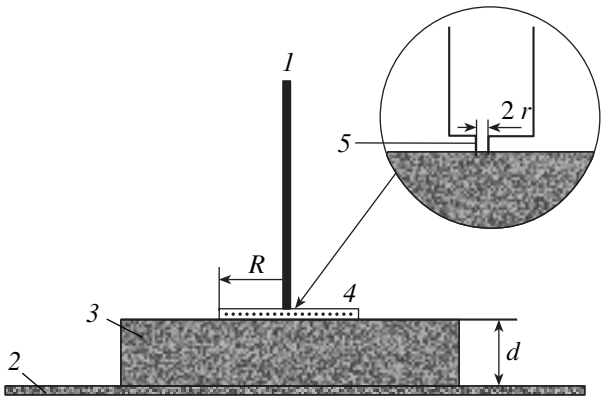


Fig. 1. Schematic disposition of electrodes for studies of the solid-dielectric breakdown: (*1*) pin electrode; (*2*) grounded plane electrode; (*3*) dielectric plate; (*4*) flashover plasma; (*5*) microprotrusion on the electrode surface. Details of the electrode contact with the solid-dielectric surface are shown in the upper insert.

Institute of Electrophysics, Ural Division,
Russian Academy of Sciences,
ul. Amundsena 106, Yekaterinburg, 620016 Russia
e-mail: mesyats@nsc.gpi.ru

1028-3358/04/4912-0727$26.00 © 2004 MAIK “Nauka/Interperiodica”

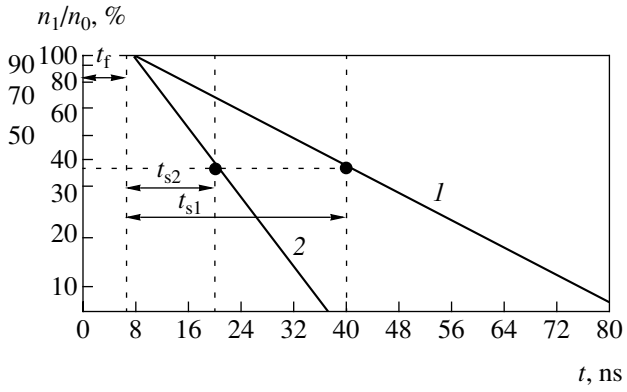


Fig. 2. Discharge-time distribution in the case of electrical breakdown of rock salt (NaCl) (d = 200 μm) for E = 2 MV cm^{-1} and the electrode area: (*1*) $S_1 = 7.1 \times 10^{-4}$ cm^2 and (*2*) $S_2 = 17.7 \times 10^{-4}$ cm^2.

the formula

$$i = \frac{dUC}{dt} = \frac{2\pi U_a \varepsilon_0 \varepsilon v^2 t}{d}. \quad (3)$$

In this case, we have taken into account that $U = U_a$ = const. If we consider this electric current to flow in metal through the microprotrusion end, then the electric current density is

$$j = \frac{i}{\pi r^2} = \frac{2U_a \varepsilon_0 \varepsilon v^2 t}{dr^2}. \quad (4)$$

It is well known the criterion for a cylindrical conductor to be electrically exploded:

$$\int_0^{t_d} j^2 dt = \bar{h}. \quad (5)$$

Here, t_d is the explosion delay time, and $\bar{h}$ is the specific action of current for the exploded conductor. For metals, the latter quantity is on the order of $\bar{h} \approx 10^9$ A^2 s cm^{-4}, whereas for graphite, it is ~10^8 A^2 s cm^{-4} [5]. We should remember that formula (5) is valid only under the condition that the heating of the microprotrusion occurs by virtue of the Joule heat release, whereas the energy loss by conduction is negligible. This takes place provided that

$$t_d < \frac{h^2 \rho c}{\lambda}, \quad (6)$$

where h is the microprotrusion height, ρ, λ, and c are, respectively, the density, heat conductivity, and heat capacity of the electrode material. For example, if $h = 10^{-4}$ cm, then this relationship is fulfilled for a number of metals and graphite as $t_d \le 10^{-8}$ s.

In accordance with formulas (4) and (5), the explosion delay time is

$$t_d = \sqrt[3]{\frac{3\bar{h}t_0^4}{4E_a^2 \varepsilon_0^2 \varepsilon^2}}, \quad (7)$$

where $E_a = \frac{U_a}{d}$ is the average value of the electric-field intensity in the solid dielectric and $t_0 = \frac{r}{v}$ is the time required for the plasma to propagate for a distance equal to the microprotrusion radius r.

Once the explosion of a microprotrusion has occurred, a thin plasma jet is formed that penetrates into the solid dielectric and initiates the formation of a discharge channel in it. Therefore, the time t_d may be interpreted as a statistical delay time t_s for the discharge. The total time t_{ed} for an electrical discharge is determined by the relation [3]

$$t_{ed} = t_s + t_f, \quad (8)$$

where t_f is the discharge formation time, i.e., the time it takes for a discharge channel to form and bridge the cathode-anode gap. It is well known that for a large number of breakdowns in solid (and also gaseous) dielectrics, the statistical distribution of the discharge delay times obeys the Laue law [3]:

$$\ln \frac{n_t}{n_0} = -\frac{1}{t_s}(t - t_f). \quad (9)$$

Here, n_0 is the total number of dielectrics broken down, and n_t is the number of samples having the discharge time t and longer. The average time t_s is determined from the condition $\frac{n_t}{n_0} = 1$.

Figure 2 shows, as an example, the statistical distribution of the discharge time for a rock-salt (NaCl) sample of thickness of 0.2 mm for an average electric-field intensity of 2×10^6 V cm^{-1} and for two different areas of graphite electrodes [3]. It follows from the given plots that the discharge formation time is t_f = 7 ns. For $S = 7.1 \times 10^{-4}$ and 17.1×10^{-4} cm^2, we have t_s = 33 and 13 ns, respectively. Hence, the average statistical breakdown delay is on the order of 10^{-8} s. Taking into account the fact that in our model, $t_d = t_s$, for the conditions of the experiment illustrated by Fig. 2, we arrive at the following values of the quantities in formula (7): $\bar{h} \approx 10^8$ A^2 s cm^{-4} (for graphite), $\varepsilon \approx 2$ (NaCl), and $E_a = 2 \times 10^6$ V cm^{-1}. In this case, in order to obtain $t_s \approx 10^{-8}$ s, it is necessary to have $t_0 \approx 10^{-11}$ s. If $r = 10^{-5}$ cm, then the velocity v must be not lower than 10^6 cm s^{-1}. This velocity is determined by the tangential component E_t of the electric-field intensity at the dielectric surface, which increases proportionally to E_a [8]. With allow-

ance for the large value of $E_a = 2 \times 10^6$ V cm^{-1}, we can find that the velocity $v = 10^6$ cm s^{-1} is quite attainable [5, 9]. Thus, in the framework of the model under consideration, the time t_s is determined by the probability of an effective microprotrusion appearance in the metal–dielectric contact zone rather than by the probability that an initiating electron will appear as was supposed in [3].

In conclusion, it is worth making several remarks.

First, the above mechanism may take place only provided that efforts for improving the electrode–dielectric contact have not been taken, i.e., as one can say, there are rims [3]. The role of these rims in the breakdown of solid dielectrics was mentioned previously in [1–4]. However, nobody before us placed emphasis upon their particular contribution to the breakdown of solid dielectrics. Rims at electrode-dielectrics contacts can be completely eliminated, e.g., in the case of the breakdown of NaCl, KCl, KBr, and KI. For these dielectrics, the water electrolyte of these salts, which fills up cylindrical depressions at two sides of the dielectric plate can serve as an electrode. However, in this case, it is worth speaking on another breakdown mechanism.

Second. The breakdown mechanism proposed by us implies that the discharge process occurring after the plasma injection into the dielectric is determined by the interaction of this plasma with the dielectric material, which proceeds in an electric field. In the case considered above (Fig. 2), the rate v_f of the discharge formation is $v_f = \dfrac{d}{t_f} \approx 3 \times 10^6$ cm s^{-1}. Analysis of the conditions governing this phenomenon needs an independent consideration.

Third, in our case, as distinct from a vacuum discharge in which the breakdown always starts at the cathode as a result of the explosion of microprotrusions under the action of the field-emission current, the solid-dielectric breakdown can occur from both the cathode and the anode. This may be explained by the fact that a dielectric flashover can take place irrespective of whether the electrode be positive or negative, which is observed experimentally [3]. The difference is related to only the values of dynamic capacities.

Finally, in the fourth place, the breakdown mechanism proposed by us can turn out to be extremely useful for the explanation of the mechanism of fracturing dielectrics and rocks by electrical discharges. In particular, this mechanism, perhaps, will make it possible to explain the effect of the discharge penetration into a solid dielectric immersed in liquid dielectric, provided that the rate of the pulsed-voltage rise exceeds a certain critical value (the so-called effect of Vorob'evs [4]).

REFERENCES

1. *Physics of Dielectrics*, Ed. by A. F. Val'ter (Gostekhteorizdat, Leningrad–Moscow, 1932).
2. S. Whitehead, *Dielectric Breakdown of Solids* (Clarendon Press, Oxford, 1951; Gosenergoizdat, Moscow, 1954).
3. A. A. Vorob'ev and G. A. Vorob'ev, *Electrical Breakdown and Destruction of Solid Dielectrics* (Vysshaya Shkola, Moscow, 1966).
4. B. V. Semkin, A. F. Usov, and V. I. Kurets, *Principles of Fracturing Materials by Electrical Pulses* (Nauka, St. Petersburg, 1995).
5. G. A. Mesyats, *Ectons in Vacuum Discharge: Breakdown, Spark, Arc* (Nauka, Moscow, 2000).
6. G. A. Mesyats, Pis'ma Zh. Éksp. Teor. Fiz. **57**, 88 (1993) [JETP Lett. **57**, 95 (1993)].
7. G. A. Mesyats, Usp. Fiz. Nauk **165**, 601 (1995) [Phys. Usp. **38**, 567 (1995)].
8. G. A. Mesyats, Dokl. Akad. Nauk **336** (5), 610 (1994) [Phys.–Dokl. **39**, 417 (1994)].
9. V. Ya. Ushakov, *The Pulsed Electrical Breakdown of Liquids* (Izd. Tomsk. Gos. Univ., Tomsk, 1975).

Translated by G. Merzon

Doklady Physics, Vol. 49, No. 12, 2004, pp. 730–731. Translated from Doklady Akademii Nauk, Vol. 399, No. 6, 2004, pp. 760–761.

PHYSICS

Work Function of Nanocrystalline Tungsten

R. R. Mulyukov* and Yu. M. Yumaguzin**

Presented by Academician R.I. Nigmatulin June 8, 2004

Received July 14, 2004

The physical properties of nanocrystalline materials with an average grain size of 10–100 nm differ significantly from the properties of coarse-grained materials [1–3]. This difference is attributable not only to the small size of nanocrystalline grains but also to changes at other structural levels of a material, including the atomic, nuclear, and electronic levels [3, 4]. In particular, we emphasize the role of outer-shell electrons in metals and alloys. The collectivization of these electrons is responsible for a large binding energy of materials and determines their properties.

In [5], qualitative and quantitative changes in the energy distributions of electrons due to the formation of the nanocrystalline structure in nickel were found by field emission spectroscopy. Two types of distributions—classical single-peak distributions and distributions with an additional maximum—were observed for different parts of the emitting surface of a point. The properties of nanocrystalline tungsten were comprehensively studied by transmission electron microscopy, field ion microscopy, field emission microscopy, and field emission spectroscopy. These studies indicated that the additional peak in the distribution of electrons over their total energies arises when the electron emission occurs from a region with a width of about 10 nm that lies in the vicinity of the grain boundaries [6]. This experiment was numerically simulated in the two-phase model [3, 4] of the nanocrystalline-material structure. This simulation supports the assumption that the formation of the nanocrystalline structure leads to the appearance of current tubes with a lower work function (grain boundaries) [7, 8]. The direct determination of the electron work function for a nanocrystalline metal is urgent from both the fundamental and applied points of view. The work function determines the operating efficiency of many types of electron sources.

* *Institute for Metal Superplasticity Problems, Russian Academy of Sciences, ul. Khalturina 39, Ufa, 450001 Russia*
e-mail: Radik@anrb.ru

** *Bashkortostan State University, ul. Frunze 32, Ufa, 450074 Russia*

In this study, the work function for electrons in a nanocrystalline metal is determined by measuring the contact potential difference. For the comparative analysis, the work function is also measured for a coarse-grained metal. We investigate 99.99%-purity tungsten. Its nanocrystalline structure was obtained by severe plastic deformation up to the true logarithmic degree $e = 7$ by torsion under quasi-hydrostatic pressure using a Bridgman anvil-type setup. The disk nanocrystalline samples were about 10 mm in diameter and 0.1–0.2 mm thick. The coarse-grained state was obtained by annealing a nanocrystalline sample for 20 min at a temperature of 1500°C. The sample microstructure was checked by means of a JEM-2000EX transmission electron microscope.

The work function was determined by electron-beam measurements of the contact potential difference (the Anderson method [9]). The measuring unit of the experimental setup is a three-electrode electron-optical system. It consists of a hot cathode, a focusing electrode, an anode with a grid hole in its central area for the ejection of a parallel electron beam into the retarding region in front of the sample, and the sample on which the external retarding potential U_r will be applied (Fig. 1). The measuring unit was placed in the work chamber of a VUP-4 vacuum setup. The measurements were performed in vacuum of $\leq 10^{-3}$ Pa. Both the measurements of the retarded current on the sample and the data processing were controlled by means of a personal computer connected with a CAMAC module. The hot cathode was manufactured from a tungsten wire. The disk samples were cut to have a diameter of about 5 mm. The retarded current being measured was amplified by a highly sensitive preamplifier and then was transmitted to an analog-to-digital converter. The retarding potential is applied to the cathode unit via a digital-to-analog converter so that the cathode–grid potential remains constant. Software employed in the measurements ensured the recording of the retarded-current intensity as a function of the retarding potential.

The system of the sample and cathode in the electric contact reaches equilibrium when their electrochemical potentials are equal to each other. In this case, the contact potential difference U_r, equal to the difference of their work functions $eU_r = \varphi - \varphi_c$, arises between them.

1028-3358/04/4912-0730$26.00 © 2004 MAIK “Nauka/Interperiodica”

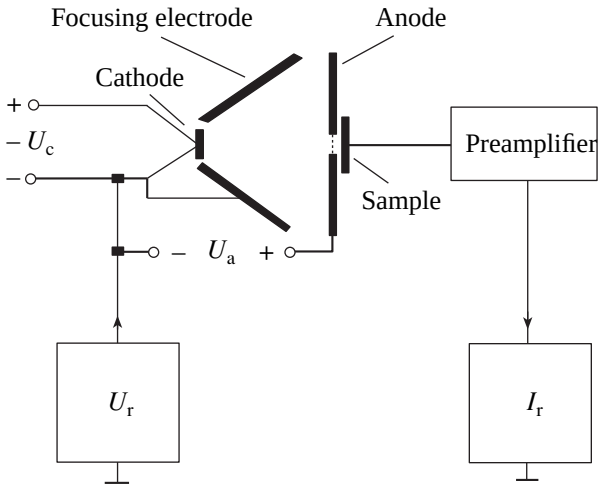


Fig. 1. Electrode system of the experimental setup and the layout of measurements.

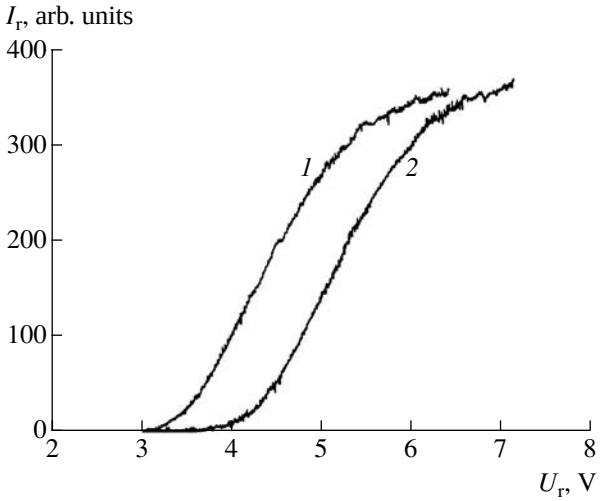


Fig. 2. Retarded-current intensity as a function of the retarding potential for (*1*) nanocrystalline and (*2*) annealed tungsten samples.

In the experiment, we measured the current I_r on the sample (retarded current) as a function of the potential difference U_r (retarding potential) between the electron source and sample. At the initial stage, the measurements were performed with a nanocrystalline tungsten sample. Later, this sample was annealed and the measurements were repeated under the same conditions. The difference $\varphi_{NC}-\varphi_0$ of the work functions for the nanocrystalline and annealed samples was determined by the relative potential shift of the $I_r(U_r)$ curves for these samples.

The studies with a transmission electron microscope showed that a nanocrystalline structure with an average grain size of about 100 nm was formed in the tungsten samples due to severe plastic deformation. The diffuse contrast at the grain boundaries and the extinction bend contours inside grains testified to the nonequilibrium state of the majority of the grain boundaries. After the annealing of nanocrystalline samples, their microstructure was recovered, and the average size of the grains in them increased to several microns.

Figure 2 shows the measured retarding curves (zero in the ordinate axis is chosen conditionally). The shift of the curves determines the difference in the tungsten work functions in the nanocrystalline and annealed coarse-grained states. The shift was found according to the intersection of the approximated linear segments of the curves with the abscissa axis. The curve for the nanocrystalline tungsten is shifted by 0.8 V towards lower potentials compared to the coarse-grained metal; i.e., the work function for the nanocrystalline metal is lower by 0.8 V.

Thus, we have found that the formation of a nanocrystalline structure results in a decrease in the electron work function of a metal. For tungsten with a grain size of about 100 nm, this decrease is equal to 0.8 V.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research, project nos. 02-01-97924-r2002agidel'_a and 03-02-16560a.

REFERENCES

1. I. D. Morokhov, L. I. Trusov, and V. I. Lapovok, *Physical Phenomena in Ultradisperse Media* (Nauka, Moscow, 1984).
2. R. Birringer and H. Gleiter, *Nanocrystalline Materials*, in *Encyclopedia of Materials. Science and Engineering. Supplement 1* (Pergamon Press, Oxford, 1988), pp. 339–351.
3. A. A. Nazarov and R. R. Mulyukov, *Handbook of Nanoscience, Engineering, and Technology* (CRC Press, Boca Raton, 2002), pp. 22-1–22-41.
4. R. R. Mulyukov and M. D. Starostenkov, Acta Met. Sin. (Eng. Lett.) **13** (1), 301 (2000).
5. L. R. Zubairov, E. A. Litvinov, R. R. Mulyukov, *et al.*, Dokl. Akad. Nauk **372**, 319 (2000) [Dokl. Phys. **45**, 198 (2000)].
6. R. R. Mulyukov, Yu. M. Yumaguzin, V. A. Ivchenko, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 377 (2000) [JETP Lett. **72**, 257 (2000)].
7. R. R. Mulyukov, E. A. Litvinov, L. R. Zubairov, *et al.*, Physica A **324** (1/4), 329 (2002).
8. E. A. Litvinov, R. R. Mulyukov, L. R. Zubairov, *et al.*, Zh. Tekh. Fiz. **74** (6), 96 (2004).
9. P. A. Anderson, Phys. Rev. **88**, 655 (1952).

Translated by G. Merzon

Doklady Physics, Vol. 49, No. 12, 2004, pp. 732–735. Translated from Doklady Akademii Nauk, Vol. 399, No. 6, 2004, pp. 762–765.

PHYSICS

Propagation of a Radio Pulse in Isotropic Plasma

G. M. Strelkov

Presented by Academician Yu. V. Gulyaev April 26, 2004

Received May 7, 2004

The problem of distortion of a radio pulse in dispersion media was posed about a century ago [1] and, as before, remains urgent by virtue of its importance for practice. In this connection, the problems of the penetration of radio pulses through the ionosphere are of independent significance. Efforts performed in this direction concern the analysis of data obtained as a result of analytical or numerical inverse Fourier transformation of the frequency spectrum of propagating signals [2–11]. The majority of the relevant studies use the approximation of a narrow-band signal. This makes it possible to significantly simplify the analysis of distortions of Gaussian or rectangular pulses under conditions of significant constraints imposed on the problem parameters. In these studies, the approximation of collisionless plasma also is usual. In the case of wide-band pulses applied in communication, radar technique, and probing of the environment, constraints of such a type turn out to be inapplicable, so that the inverse Fourier transformation can be performed only numerically (see, e.g., [12, 13]).

In the present paper, we develop an approach capable of describing the propagation of radio pulses in an ionized medium with allowance made for collisions. The approach is based on the time representation of a radio signal. The analytical description for the space-time evolution of the initial envelope of a sufficiently general form is obtained.

The propagation of a radio pulse is described by the wave equation

$$\frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} + \frac{4\pi}{c^2}\frac{\partial^2 P}{\partial t^2}, \quad (1)$$

where E is the electromagnetic-field intensity, c is the speed of light, z is the pulse-propagation direction, t is time, and P is the polarization of a unit volume of the medium. According to the model of a medium with free charges [14, 15], the polarization P is described by the equation

$$\frac{\partial^2 P}{\partial t^2} + \nu_e \frac{\partial P}{\partial t} = \frac{e^2 N}{m} E. \quad (2)$$

In Eq. (2), e and m are the electron charge and mass, respectively; N is the electron concentration; and ν_e is the effective collision frequency, which takes into account the electron energy loss in collisions with neutral molecules and ions.

At the boundary of the half-space $z \geq 0$, in which a pulse propagates, we can write out the field E in the form

$$E(0; t) = A(0; t)\exp(i\omega t), \quad t \geq 0. \quad (3)$$

Here, $\omega = 2\pi f$, f is the carrying frequency (or the filling frequency), and $A(0; t)$ is the pulse envelope for $z = 0$.

The pulse leading front always propagates at the velocity of light in the given medium. Correspondingly, the field E is sought in the form

$$E(z; t) = \begin{cases} A\left(z; t - \frac{z}{c}\right)\exp(i(\omega t - kz)), & t - \frac{z}{c} \geq 0, \\ 0, & t - \frac{z}{c} < 0, \end{cases} \quad (4)$$

where $k = \frac{\omega}{c} = \frac{2\pi}{\lambda}$ is the wave vector.

We now replace the variables in Eqs. (1) and (2):

$$z' = z; \quad t' = t - \frac{z}{c}. \quad (5)$$

Then, with allowance made for expressions (4), we obtain

$$\frac{\partial^2 A}{\partial (z')^2} - \frac{2}{c}\frac{\partial^2 A}{\partial z' \partial t'} - 2ik\frac{\partial A}{\partial z'} = \frac{4\pi}{c^2}\frac{\partial^2 P}{\partial (t')^2}\exp(-i\omega t'), \quad (6)$$

$$\frac{\partial^2 P}{\partial (t')^2} + \nu_e \frac{\partial P}{\partial t'} = \frac{e^2 N}{m} A(z'; t')\exp(i\omega t'). \quad (7)$$

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, pl. Akademika Vvedenskogo 1, Fryazino, Moscow oblast, 141190 Russia
e-mail: strelkov@ms.ire.rsst.ru

1028-3358/04/4912-0732$26.00 © 2004 MAIK "Nauka/Interperiodica"

We compare the first and third terms in the left-hand side of Eq. (6) as applied to the radio pulse (with filling) under consideration. Its characteristic duration t_p obeys the inequality $ft_p \gg 1$, whereas the pulse, in itself, occupies the interval $L_p = ct_p \gg \lambda$ in the z' axis. Thus, the estimate

$$\left|\frac{\partial^2 A}{\partial (z')^2}\right| \sim \left|\frac{1}{L_p}\frac{\partial A}{\partial z'}\right| \ll \left|\frac{4\pi}{\lambda}\frac{\partial A}{\partial z'}\right| \quad (8)$$

is valid and the first term in the left-hand side of Eq. (6) can be ignored.

As the position and velocity of an electron cannot vary instantaneously, at the moment of the pulse arrival at the point z', the conditions

$$P(z'; 0) = 0, \quad \left.\frac{\partial P}{\partial t'}\right|_{t' = 0} = 0$$

are fulfilled. Then, the solution to Eq. (7) is

$$P(z'; t') = \frac{e^2 N}{m}\int_0^{t'} E(z'; \theta)\frac{1 - \exp(-\nu_e(t' - \theta))}{\nu_e} d\theta. \quad (9)$$

We now substitute solution (9) into Eq. (6). With due regard for estimate (8), the equation for the pulse envelope $A(z'; t')$ takes the form

$$\frac{\partial^2 A}{\partial z' \partial t'} + i\omega\frac{\partial A}{\partial z'} = -\frac{2\pi e^2 N}{mc}A(z'; t') + \frac{2\pi e^2 N}{mc}\nu_e\int_0^{t'} A(z'; \theta)\exp(-(\nu_e + i\omega)(t' - \theta))d\theta. \quad (10)$$

Furthermore, using the Laplace transforms in terms of the variable t' to expression (10), we arrive at

$$F(z'; p) = F(p)\exp\left(-\frac{\delta}{p + \nu_e + i\omega}\right). \quad (11)$$

Here, the notation $F(z'; p) = \int_0^\infty A(z'; t)\exp(-pt')dt'$; $F(p) = \int_0^\infty A(0; t')\exp(-pt')dt'$ is the Laplacian image of the initial envelope; $\delta = \delta(z) = \frac{\tau(z')}{2}\frac{\omega^2}{\nu_e}$; $\tau = \gamma z'$ is the route optical depth; $\gamma = \frac{4\pi e^2 N}{m\omega^2}\frac{\nu_e}{c}$ is the power absorption coefficient of plasma [6] (for $\nu_e \ll \omega$); and it is assumed that $A(z'; 0) = 0$. The last condition is evidently fulfilled for a pulse whose initial envelope is zero at $t = 0$. Performing the inverse Laplace transform, we find

$$A(z'; t') = A(0; t') - \int_0^{t'}\frac{\sqrt{\delta}}{\sqrt{t' - \theta}}J_1(2\sqrt{\delta(t' - \theta)}) \times \exp(-(\nu_e + i\omega)(t' - \theta))A(0; \theta)d\theta \quad (12)$$

[here and below, $J_k(x)$ is the Bessel function].

Based on expression (12), we consider distortions of a biexponential pulse with its envelope,

$$A(0; t) = A_0\left(\exp\left(-\frac{\alpha t}{t_p}\right) - \exp\left(-\frac{\beta t}{t_p}\right)\right). \quad (13)$$

This pulse is a good approximation for those with different slopes of the leading and trailing fronts. Substituting (13) into (12), we have

$$A(z'; t') = A(z'; t'; \alpha) - A(z'; t'; \beta), \quad (14)$$

where

$$A(z'; t'; \alpha) = A_0\exp\left(-\frac{\alpha t'}{t_p}\right) \times\left(1 - \int_0^{t'}\frac{\sqrt{\delta}}{\sqrt{\theta}}J_1(2\sqrt{\delta\theta})\exp\left(-\left(\nu_e - \frac{\alpha}{t_p} + i\omega\right)\theta\right)d\theta\right). \quad (15)$$

Performing integration in (15) by parts an infinitely large number of times, we reduce the expression $A(z'; t'; \alpha)$ to the form

$$A(z'; t'; \alpha) = A_0\exp\left(-\frac{\alpha t'}{t_p}\right)\left(1 - \sum_{k=1}^{\infty}(-1)^{k+1}\left(\frac{\delta}{\nu_e - \frac{\alpha}{t_p} + i\omega}\right)^k - \sum_{k=1}^{\infty}(-1)^{k+1}\left(\frac{\delta}{\nu_e - \frac{\alpha}{t_p} + i\omega}\right)^k\frac{J_k(2\sqrt{\delta t'})}{(\sqrt{t'})^k} \times\exp\left(-\left(\nu_e - \frac{\alpha}{t_p} + i\omega\right)t'\right)\right). \quad (16)$$

Summing the two series (the second of them is the Neu-

mann series), we obtain

$$A(z';\,t';\,\alpha) = A_0\exp\left(-\frac{\alpha t'}{t_p}\right)\left\{\exp\left(-\frac{\delta}{\nu_e-\frac{\alpha}{t_p}+i\omega}\right)\right.$$

$$+\exp\left(-\left(\nu_e-\frac{\alpha}{t_p}+i\omega\right)t'\right)\left[\frac{\sqrt{\delta}}{\left(\nu_e-\frac{\alpha}{t_p}+i\omega\right)\sqrt{t'}}J_1(2\sqrt{\delta t'})\right.$$

$$-\frac{1}{4\left(\nu_e-\frac{\alpha}{t_p}+i\omega\right)(t')^2}\exp\left(-\frac{\delta}{\nu_e-\frac{\alpha}{t_p}+i\omega}\right) \tag{17}$$

$$\left.\left.\times\int_0^{2\sqrt{\delta t'}}dy y^2 J_1(y)\exp\left(\frac{y^2}{4\left(\nu_e-\frac{\alpha}{t_p}+i\omega\right)}\right)\right]\right\}.$$

Finally, performing integration by parts in (17) and introducing a new variable $x=\dfrac{y}{2\sqrt{\delta t'}}$, we arrive at the final expression for the first component for the biexponential-pulse envelope [see (14)]:

$$A(z';\,t';\,\alpha) = A_0\exp\left(-\frac{\alpha t'}{t_p}\right)\exp\left(-\frac{\delta}{\nu_e-\frac{\alpha}{t_p}+i\omega}\right)$$

$$\times\left[1+\exp\left(-\left(\nu_e-\frac{\alpha}{t_p}+i\omega\right)t'\right)\frac{2\delta}{\nu_e-\frac{\alpha}{t_p}+i\omega}\right. \tag{18}$$

$$\left.\times\int_0^1 xJ_0(2x\sqrt{\delta t'})\exp\frac{\delta x^2}{\nu_e-\frac{\alpha}{t_p}+i\omega}dx\right].$$

The term $A(z';\,t';\,\beta)$ is obtained from $A(z';\,t';\,\alpha)$ by substituting α for the quantity β.

Distortions of a so-called sine pulse are analyzed in a similar manner. The initial pulse envelope is given by the expression

$$A(0;\,t) = \begin{cases}\sin\dfrac{\pi t}{t_p}, & t\le t_p,\\ 0, & t>t_p.\end{cases} \tag{19}$$

We represent envelope (19) in the form

$$A(0;\,t) = \begin{cases} f(0;\,t), & t\le t_p,\\ f(0;\,t)+f(0;\,t-t_p), & t>t_p,\end{cases} \tag{20}$$

where

$$f(0;\,t) = \frac{1}{2i}\left(\exp i\pi\frac{t}{t_p}-\exp\left(-i\pi\frac{t}{t_p}\right)\right),$$

and substitute (20) into formula (12). Performing operations analogous to those described above, we arrive at the following expressions for the deformed sine pulse:

$$A(z';\,t') = f(z';\,t') = \frac{\exp\frac{i\pi t'}{t_p}}{2i}\exp\frac{\delta}{\nu_e+\frac{i\pi}{t_p}+i\omega}$$

$$\times\left[1+\exp\left(-\left(\nu_e+\frac{i\pi}{t_p}+i\omega\right)t'\right)\frac{2\delta}{\nu_e+\frac{i\pi}{t_p}+i\omega}\right.$$

$$\left.\times\int_0^1 xJ_0(2x\sqrt{\delta t'})\exp\frac{\delta x^2}{\nu_e+\frac{i\pi}{t_p}+i\omega}dx\right]$$

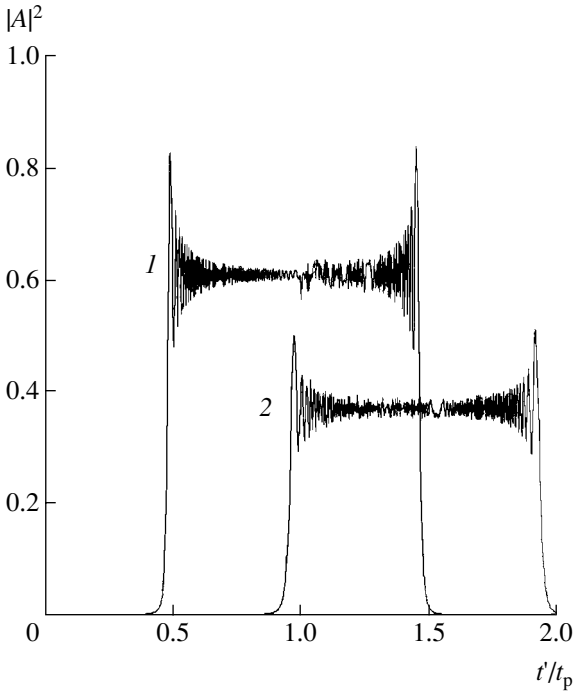
$$-\frac{\exp\left(-\frac{i\pi t'}{t_p}\right)}{2i}\exp\frac{\delta}{\nu_e-\frac{i\pi}{t_p}+i\omega} \tag{21}$$

$$\times\left[1+\exp\left(-\left(\nu_e-\frac{i\pi}{t_p}+i\omega\right)t'\right)\frac{2\delta}{\nu_e-\frac{i\pi}{t_p}+i\omega}\right.$$

$$\left.\times\int_0^1 xJ_0(2x\sqrt{\delta t'})\exp\frac{\delta x^2}{\nu_e-\frac{i\pi}{t_p}+i\omega}dx\right],\quad t'\le t_p;$$

$$A(z';\,t') = f(z';\,t')+f(z';\,t'-t_p),\quad t'>t_p. \tag{22}$$

We now obtain the solution to the problem for the rectangular pulse shape. In this connection, we note that, in the case of the simultaneously tending $\alpha\longrightarrow 0$ and $\beta\longrightarrow\infty$, a biexponential pulse transforms into a step-function signal with the amplitude A_0, the zero value of A being conserved as $t=0$. Using the corresponding



Deformation of the envelope of a rectilinear radio pulse in homogeneous isotropic plasma for the route optical depth (*1*) $\tau = 0.5$ and (*2*) 1.0.

transformations in (14) and (18) at $t' > 0$, we find the envelope of the initial step-function signal in the form

$$A(z';\, t) = A_0 \exp\left(-\frac{\delta}{\nu_e + i\omega}\right)\left[1 + \exp(-(\nu_e + i\omega)t') \times \frac{2\delta}{\nu_e + i\omega}\int_0^1 x J_0(2x\sqrt{\delta t'}) \exp\frac{\delta x^2}{\nu_e + i\omega} dx\right]. \quad (23)$$

The passage from the envelope of a step-function signal (23) to the envelope of a rectangular pulse occurs by analogy with formula (22). The figure illustrates features of the envelope deformation of a rectangular pulse for the parameter values $f = 2.8$ MHz, $t_p = 0.5$ ms, and $\nu_e = 10^3\ s^{-1}$ (see 12, 13]). The pulse weakens and delays proportionally to the route optical depth as compared to the hypothetical case of its propagation at the speed of light.

Finally, as $t' \longrightarrow \infty$, we obtain from (23)

$$A(z') = A_0 \exp\left(-\frac{\delta}{\nu_e + i\omega}\right) = A_0 \exp\left(-\frac{\frac{\tau}{2}}{1 + \left(\frac{\nu_e}{\omega}\right)^2}\right) \exp\left(-i\left(\frac{\frac{4\pi e^2 N}{m}\frac{z'}{2\omega c}}{1 + \left(\frac{\nu_e}{\omega}\right)^2}\right)\right). \quad (24)$$

The first multiplier in (24) describes a decrease in the amplitude, whereas the second one corresponds to the phase increment of the plane wave when it passes by the path z' [6]. This was expected, since, with increasing t', the problem of the incidence of a step-function signal onto the semi-infinite homogeneous medium transforms into that of the propagation of a plane wave in a homogeneous medium.

REFERENCES

1. A. Sommerfeld, Ann. Phys. (New York) **44** (10), 177 (1914).
2. L. A. Zhekulin, Izv. Akad. Nauk SSSR, Otd. Tekh. Nauk, No. 7, 91 (1938).
3. N. G. Denisov, Zh. Éksp. Teor. Fiz. **21**, 1354 (1951).
4. B. N. Gershman, Zh. Tekh. Fiz. **22**, 101 (1952).
5. J. R. Wait, Radio Sci. **69D**, 1387 (1965).
6. V. L. Ginzburg, *Propagation of Electromagnetic Waves in Plasmas* (Fizmatgiz, Moscow, 1967).
7. G. I. Terina, Radiotekh. Élektron. (Moscow) **12**, 123 (1967).
8. G. I. Terina, Radiotekh. Élektron. (Moscow) **17**, 611 (1972).
9. L. A. Vaĭnshteĭn, Usp. Fiz. Nauk **118**, 339 (1976).
10. G. M. Whitman and L. B. Felsen, IEEE Trans. Ant. Propagat. **AP-28** (3), 342 (1980).
11. N. V. Kretov, T. E. Ryzhkina, and L. V. Fedorova, Radiotekh. Élektron. (Moscow) **36**, 1 (1991).
12. G. M. Strelkov, O. S. Gashevskaya, and O. G. Derkach, in *Reports of 9th Regional Conference on Radio Wave Propagation* (NIIRF, St. Petersburg, 2003), p. 36.
13. A. I. Rogashkova, in *Reports of 9th Regional Conference on Radio Wave Propagation* (NIIRF, St. Petersburg, 2003), p. 50.
14. M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *Theory of Waves* (Nauka, Moscow, 1979).
15. E. A. Pamyatnykh and E. A. Turov, *Principles of Electrodynamics of Material Media for Time-Dependent and Inhomogeneous Media* (Nauka, Moscow, 2000).

Translated by G. Merzon

Doklady Physics, Vol. 49, No. 12, 2004, pp. 736–738. Translated from Doklady Akademii Nauk, Vol. 399, No. 5, 2004, pp. 617–619.

TECHNICAL PHYSICS

Singular Integral Representation of the Electromagnetic Field of an Electric Dipole in the Near-Field Zone

V. A. Neganov

Presented by Academician Yu.V. Gulyaev June 7, 2004

Received June 24, 2004

ESSENCE OF THE PROBLEM

The problem of the determination of electromagnetic fields in the immediate vicinity of radio devices (i.e., from the standpoint of the antenna-theory terminology, in the antenna's near-field zone) is associated with studies of electromagnetic compatibility, electromagnetic ecology, and methods of antenna measurements. Usually, the electromagnetic field generated by an electric dipole (Fig. 1) is calculated on the basis of the z component of the electrodynamic vector potential A_z^e for the electric current. This component is determined in terms of the component of the electric current $I_z(z) = 2\pi a \eta_z^e(z)$ on the dipole (η_z^e is the component of the surface-current density on the dipole and a is the dipole radius) [1–3]:

$$A_z^e(\rho, z) = \int_{-l}^{l} I_z(z')G(\rho, z-z')dz, \qquad (1)$$

where

$$G(\rho, z-z') = \frac{1}{4\pi R}\exp\{-ikR\}. \qquad (2)$$

Here, $R = \sqrt{(z-z')^2 + \rho^2}$, $k = \dfrac{\omega}{c\sqrt{\varepsilon\mu}}$, ε and μ are the permittivity and permeability of the dipole environment, respectively, and l is the dipole length. It is evident that $G(\rho, z-z')$ is the free-space Green's function of a point source located at the point $\{\rho = 0, z = z'\}$. The unknown electric-current distribution $I_z(z)$ on the dipole is usually found from either the Pocklington or Hallen integral equations [1–3]. Based on the function $I_z(z)$ and employing the usual differentiation of expression (1) with respect to the coordinates ρ and z [1–3], it is easy to derive expressions for the components of the electromagnetic field generated by the dipole at any point in space.

The reliability of the fields **E** and **H** numerically calculated for the near-field zone of the electric dipole must be verified for at least two reasons. First, the determination of the unknown electric current $I_z(z)$ on the dipole from the Pocklington and Hallen integral equations (Fredholm integral equations of the first kind) leads to an ill-posed problem (see [4]). Second, the use of Green's function (2) to calculate the field leads to a self-inconsistent setting of the problem, because the passage to the limit from the near-field zone to the field (current) on the dipole surface is absent. By the method of singular integral equations developed in [5–7], the problem of calculating the electric current on the dipole is reduced to the Fredholm integral equation of the second kind. This approach provides for a mathematically correct determination of the surface-current density on the dipole.

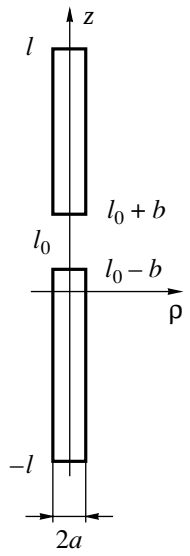


Fig. 1. Electric dipole geometry.

Volga State Academy of Telecommunications and Informatics,
ul. L'va Tolstogo 23, Samara, 443010 Russia
e-mail: neganov_samara@psati.ru

1028-3358/04/4912-0736$26.00

In this study, which generalizes the results of [5–7], an algorithm is described for calculating the electromagnetic field of a radiating structure at an arbitrary point in space. The singular integral representations obtained for the electromagnetic-field components make it possible to calculate the field at any point in space by integrating the longitudinal component of the surface-current density on the dipole. One of the important advantages of these relationships is that, on the dipole surface, they take the form of a singular integral equation [7] for determining the unknown surface-current distribution. Then, this distribution can be used to find the electromagnetic field of the dipole. It is worth noting that the method of the singular integral representation of the electromagnetic field was apparently proposed for the first time for solving the inner problems of natural waves of shielded stripe-line microwave structures [8].

SETTING OF THE PROBLEM. SINGULAR INTEGRAL REPRESENTATION FOR FIELD COMPONENTS

We consider the electromagnetic field of a perfectly conducting tubular electric dipole [7] of length $2l$ and radius a, which is excited in the gap region ($z \in [l_0 - b, l_0 + b]$) by a high-frequency generator (Fig. 1). Under the assumption that the field is independent of the φ coordinate, Maxwell's equations split into two independent sets with respect to the components $\{E_\rho, E_z, H_\varphi\}$ and $\{E_\varphi, H_\rho, H_z\}$. We start from the set of Maxwell's equations that describes the behavior of the components E_ρ, E_z, and H_φ. In this case, only the longitudinal component of the surface-current density exists on the dipole surface.

The initial expression required to obtain the singular integral representation of the electromagnetic field is relationship (1) for the z component of the electrodynamic vector potential A_z^e for the electric current expressed in terms of the electric current $I_z(z)$ on the dipole, but with the other Green's function [9]

$$G(\rho, z - z') = \frac{1}{8\pi i}\int_{-\infty}^{\infty} e^{-ih(z-z')} J_0(-ia\nu) H_0^{(2)}(-i\rho\nu)\,dh. \quad (3)$$

Here, $\nu = \sqrt{h^2 - k^2}$, $J_0(x)$ is the zero-order Bessel function of the first kind, and $H_0^{(2)}(x)$ is the zero-order Hankel function of the second kind. It is easy to show that Green's function (3) has a logarithmic singularity as $z \longrightarrow z'$.

Using potential (1) to separate the singularity in Eq. (3) in explicit form, we can write the singular integral representations

$$E_\rho = \frac{1}{i\omega\varepsilon_0\varepsilon}\int_{-l}^{l} J(z')[G_\rho(\rho, z - z') + S_1(\rho, z - z')]\,dz',$$

$$E_z = \frac{1}{i\omega\varepsilon_0\varepsilon}\int_{-l}^{l} J(z')[G_z(\rho, z - z') + S_2(\rho, z - z')]\,dz', \quad (4)$$

$$H_\varphi = -\int_{-l}^{l} I_z(z')[G_\rho(\rho, z - z') + S_1(\rho, z - z')]\,dz'.$$

These representations determine the electromagnetic field of the electric dipole at any point in space in terms of both the current $I_z = 2\pi a \eta_z^e$ and its derivative $J = \frac{dI_z}{dz}$ with respect to the coordinate z. The Green's functions G_ρ and G_z belong to the space of square integrable functions and are integrals whose integrands decrease rapidly as their arguments increase. The functions S_1 and S_2 have the isolated singularities

$$S_1(\rho, z - z') = -\frac{1}{4\pi^2\sqrt{a\rho}}\left[\frac{\rho - a}{(z - z')^2 + (\rho - a)^2}\right],$$
$$S_2(\rho, z - z') = -\frac{1}{4\pi^2\sqrt{a\rho}}\left[\frac{z - z'}{(z - z')^2 + (\rho - a)^2}\right]. \quad (5)$$

For $\rho = a$, $z \in [-l, l]$ (i.e., for the field on the dipole surface), we arrive at the expressions

$$S_1 = -\frac{1}{4\pi a}\delta(z - z'),$$
$$S_2(a, z - z') = -\frac{1}{4\pi^2 a(z - z')},$$

where $\delta(z - z')$ is the delta function. In this case, expressions (4) yield the values of the electromagnetic-field components on the dipole surface.

We now use the boundary conditions on the surface $\rho = a$:

$$E_z = \begin{cases} 0 & \text{for } z \in [-l, l_0 - b] \cup [l_0 + b, l], \\ E_z^{\text{ex}} & \text{for } z \in [l_0 - b, l_0 + b], \end{cases}$$

where E_z^{ex} is the z component of the extraneous electric field in the dipole gap. In this case, instead of the second relation in (4), we can write the following singular

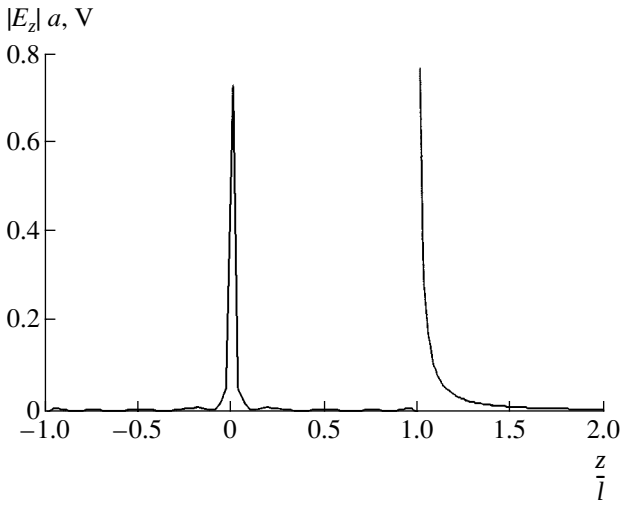


Fig. 2. Product $|E_z|a$ vs. the normalized coordinate $\frac{z}{l}$ for $\rho = a$.

integral equation for the determination of the unknown function $J(t)$ ($t \in [-1, 1]$):

$$\frac{1}{\pi}\int_{-1}^{1}\frac{J(t')}{t-t'}dt' + \int_{-1}^{1}J(t')M(t-t')dt = -i\sigma E_z^{\text{ex}}(t), \quad (6)$$

where $M(t-t')$ is the known regular kernel and σ is the known constant. The dimensionless variables $t = \frac{z}{l}$ and $t' = \frac{z'}{l}$ are used in Eq. (6). We can see that Eq. (6) coincides with the singular integral equation obtained in [7] by another method.

ELECTROMAGNETIC FIELD OF THE ELECTRIC DIPOLE IN THE NEAR-FIELD ZONE. NUMERICAL RESULTS

Singular integral representations (4) of the dipole electromagnetic field, together with singular integral equation (6), allow us to mathematically justify an approach to calculating the electromagnetic field of the electric dipole in the near-field zone. As an example, Fig. 2 shows the distribution obtained by this method for the modulus of the component E_z of the electric field for dipole radiation on the surface $\rho = a$ as a function of the coordinate $t = \frac{z}{l}$. The calculation was performed for the symmetric dipole (in Fig. 1, $l_0 = 0$) with the geometric parameters $\frac{l}{\lambda} = \frac{1}{4}$, $\frac{a}{\lambda} = \frac{1}{400}$, and $\frac{b}{l} = \frac{1}{100}$ and a voltage of 1 V across the gap. As is seen, for $t \in [-1, 1]$ (on the surface of the electric dipole), the z component vanishes on the metallic surface; i.e., the boundary condition is fulfilled. In the gap $\left(t \in \left[-\frac{b}{l}, \frac{b}{l}\right]\right)$, the modulus of E_z is equal to the extraneous exciting function E_z^{ex}. For $|t| > 1$ and near the dipole end walls, the E_z component tends to infinity, which corresponds to the behavior of the electromagnetic field in the vicinity of metallic ribs [10].

The singular integral representation obtained in this work for the electromagnetic field provides the transition to the limit from the electromagnetic field in the near-field zone of the electric dipole to the field on the dipole surface. For comparison, the commonly used approach to calculating the electromagnetic field on the basis of Green's function (2) does not ensure this possibility. In particular, the electromagnetic field obtained in the latter approach near metallic ribs does not satisfy the condition on the rib [10]. The approach proposed in our paper can be generalized for other radiating structures.

REFERENCES

1. D. M. Sazonov, *Antennas and Microwave Devices. Textbook for Radio Engineers* (Vysshaya Shkola, Moscow, 1988).
2. D. I. Voskresenskiĭ, V. L. Gostyukhin, V. M. Maksimov, and L. I. Ponomarev, *Antennas and Microwave Devices* (Izd. MAI, Moscow, 1999).
3. *Computer Techniques for Electromagnetics*, Ed. by R. Mittra (Pergamon, Oxford, 1973; Mir, Moscow, 1977).
4. A. N. Tikhonov and V. Ya. Arsenin, *Methods for Solving Ill-Posed Problems* (Nauka, Moscow, 1986; Halsted Press, New York, 1977).
5. V. A. Neganov, I. V. Matveev, and S. V. Medvedev, Pis'ma Zh. Tekh. Fiz. **26** (12), 86 (2000) [Tech. Phys. Lett. **26**, 535 (2000)].
6. V. A. Neganov and I. V. Matveev, Dokl. Akad. Nauk **371**, 36 (2000) [Dokl. Phys. **45**, 317 (2000)].
7. V. A. Neganov, E. I. Nefedov, and G. P. Yarovoĭ, *Electrodynamic Methods for Designing Microwave Devices and Antennas*, Ed. by V. A. Neganov (Radio i Svyaz', Moscow, 2002).
8. V. A. Neganov, Radiotekh. Élektron. (Moscow) **34**, 2251 (1989).
9. G. T. Markov and A. F. Chaplin, *Excitation of Electromagnetic Waves* (Énergiya, Moscow, 1967).
10. R. Mittra and S. W. Lee, *Analytical Techniques in the Theory of Guided Waves* (Macmillan, New York, 1971; Mir, Moscow, 1974).

Translated by G. Merzon

Doklady Physics, Vol. 49, No. 12, 2004, pp. 739–742. Translated from Doklady Akademii Nauk, Vol. 399, No. 6, 2004, pp. 766–769.

TECHNICAL PHYSICS

Vacuum Deposition of Thin Teflon-Like Films from a Supersonic C_2F_4 Jet

Academician A. K. Rebrov, N. I. Timoshenko, and A. V. Shishkin

Received September 3, 2004

Teflon-like films are produced by deposition from a gas-phase of various fluorocarbons and their mixtures or fragments of polytetrafluoroethylene (PTFE) with various-size molecular chains (including monomers). The flow composition and energy states of particles being deposited are specified by the technology employed, whereas the properties of the film to be produced essentially depend on the nature, state, and temperature of the substrate. To produce teflon-like films by deposition from a gas-phase, the following methods are most often used: ion-beam deposition [1], electron-beam [2] and laser-beam [3] ablations, plasma chemical processes [4], thermal destruction of PTFE with evaporation [5], and thermal activation of gases containing fluorocarbons [6]. The common feature typical of the methods listed above is the indeterminacy of the concentration and molecular composition of the gas phase produced. In this paper, we studied the deposition of thin teflon-like films from a free supersonic jet consisting of products of PTFE thermal decomposition (including the monomeric state) in a reactor. The use of such a "gasification" of PTFE has a number of advantages.

From the published data [7, 8] it is known that, at a pressure of several Torr and at a temperature exceeding 800 K (in the case of thermodynamic equilibrium), the thermal decomposition of PTFE results in a 97% yield of C_2F_4. The gas-dynamic and kinetic properties of C_2F_4 are well studied and can be employed to estimate the flow characteristics.

In the method that we used, high-temperature decomposition products of a polymer loaded into a reactor in the condensed state or delivered to it in the gas phase are expanded into a vacuum chamber through a sonic or supersonic nozzle. Upon expanding the flow from the sonic nozzle into immobile gas with a low pressure P_l, a jet with a barrel-shaped shock wave is formed. The jet is terminated by the Mach disk located at a known distance, namely, $l_M = 0.67d\left(\frac{P_0}{P_l}\right)^{0.5}$ from the nozzle. Here, P_0 is the stagnation pressure and d is the nozzle diameter. Based on the recommendations of [9], we have performed the majority of our experiments in a configuration in which the substrate was situated in front of the Mach disk with $P_0 d \approx 50$ Torr mm and $\frac{P_0 d}{l} \approx 10$ Torr, where l is the nozzle–substrate spacing.

Under the above conditions and at the stagnation temperature $T_0 \approx 900$ K, the maximum flow velocity was higher than 100 m s^{-1}. In the experiments, a compressed layer joined to the shock wave was formed in front of the substrate.

The experiments were carried out using a gas-dynamic low-density setup belonging to the Institute of Thermal Physics, Siberian Division, Russian Academy of Sciences. It was possible to control the deposition process and, as a consequence, to modify the structure and properties of the film produced by varying different parameters such as T_0, P_0, P_l, d, l, the nozzle shape, the structure of the substrate surface and its temperature T_s, the geometric orientation of a substrate with respect to the jet, or by using a carrier gas with a different nature and relative concentration, as well as an external energy action on the flow. Using a set of substrate holders, we investigated the effect of different materials on the structure of the film produced. For materials we used the following: (a) polished surfaces of metals and alloys (copper, carbon steel, stainless steel, and an aluminum alloy); (b) surfaces of silicon, vacuum rubber, organic glass, PTFE, nylon, and vacuum-oil films. The basic parameters of the deposition process were the following: $P_0 = 5$ Torr, $T_0 = 620$°C, $P_l = 4 \times 10^{-2}$ Torr, and $T_s = 50$°C. In this and other cases, the morphology of the films was examined with an LEO-420 scanning electron microscope. The surface IR spectra were measured with an IFS-66 Fourier spectrometer. For the above-indicated parameters of the experiment, different-scale dendritic structures that formed a developed surface were observed. The surface area measured by

Kutateladze Institute of Thermal Physics, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 1, Novosibirsk, 630090 Russia

e-mail: rebrov@itp.nsc.ru; vika@itp.nsc.ru; andrshi@itp.nsc.ru

1028-3358/04/4912-0739$26.00

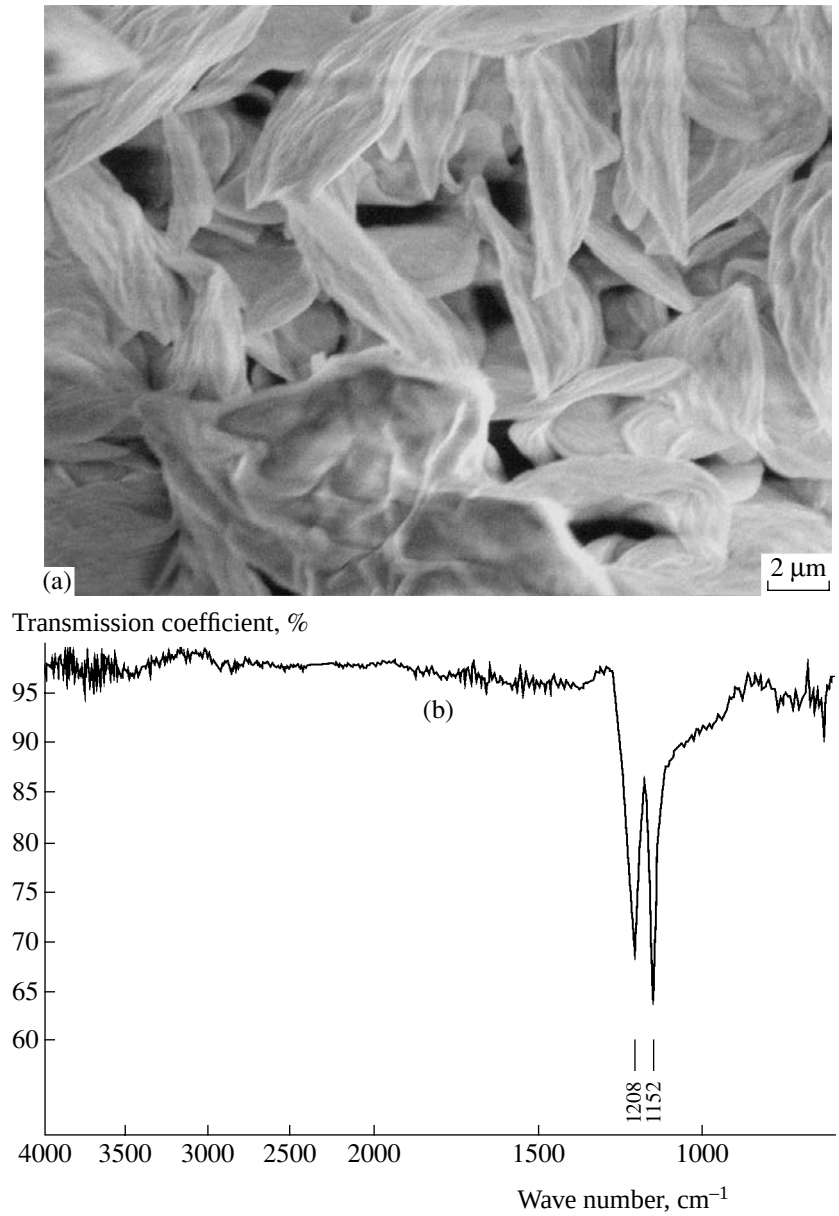


Fig. 1. (a) Image of a film deposited at a temperature of $T_s \approx 50$–150°C for the helium carrier-gas mass fraction $y_{He} \approx 50\%$ and (b) the Fourier transform of the film IR absorption spectrum.

the method of gas sorption has shown its enlargement by a factor of ~1500 compared to the geometric projection. Depending on the type of substrate, the size of the primary branches d_1 varied from 2 to 10 μm. It was found that the value of d_1 and the film porosity correlate with the thermal diffusivity of the substrate material and its chemical composition.

The effect of T_s on the properties of the films was investigated by simultaneously depositing films onto four polished copper plates (10 mm in diameter) that had different surface temperatures as measured by thermocouples. In accordance with the estimate of the equilibrium expansion of C_2F_4 (with the adiabatic index $\gamma \approx 1.1$), the jet parameters at the substrate location point are the following: the Mach number is $M \approx 4$, the Knudsen number is $Kn \approx 0.5$, and the temperature is $T_f \approx 230$°C.

Similar experiments were performed with helium as a carrier gas. The mass fraction of He in the reactor was $y_{He} \approx 50\%$. In this case, the average relative molecular mass of the mixture and the adiabatic index γ were 55 and ≈1.32, respectively, whereas the temperature of the expanding mixture near the substrate was $T_f \approx 130$°C. When the carrier gas was used, a much more intense deposition was observed.

As an example, we present here the most contrasting structures obtained in this study. Figure 1a shows the

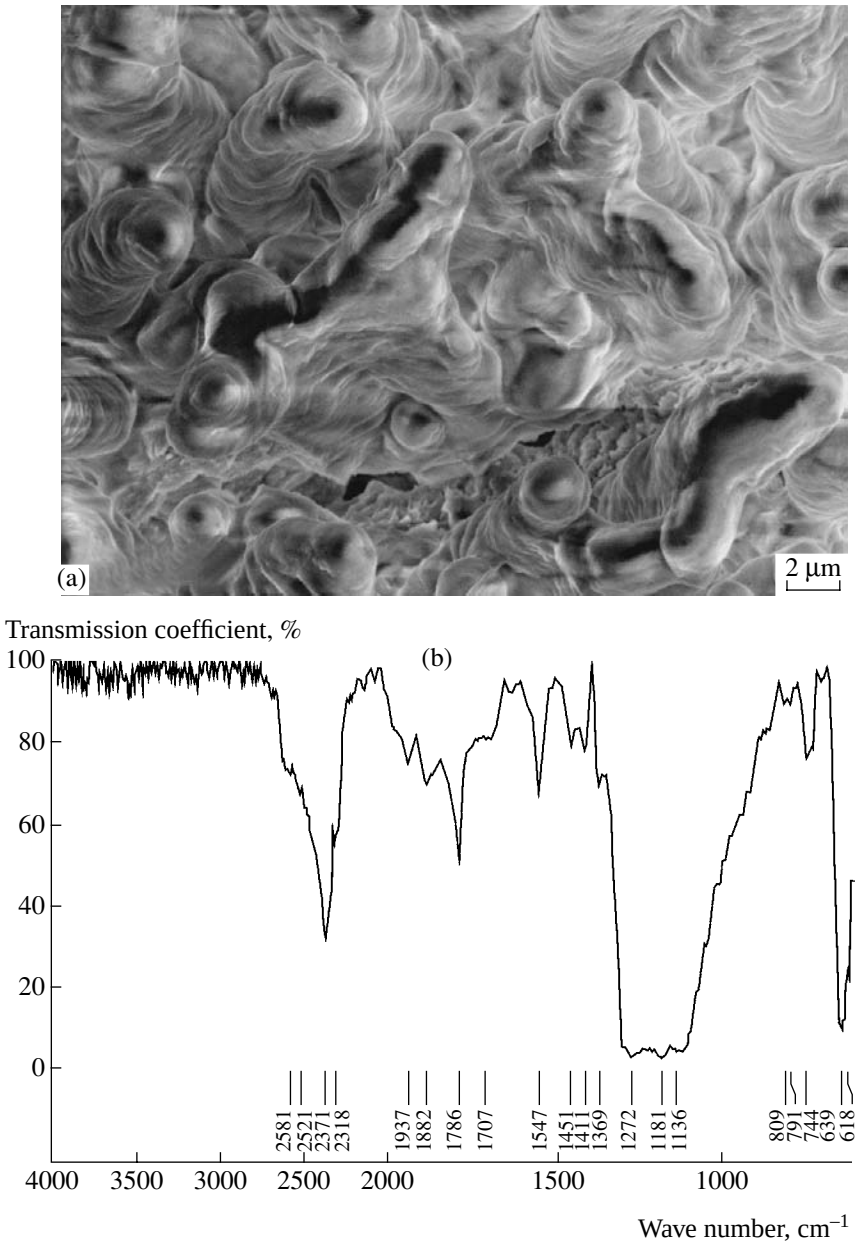


Fig. 2. (a) Image of a film deposited at a temperature of $T_s \approx 290$°C for the helium carrier-gas mass fraction $y_{He} \approx 0\%$ and (b) the Fourier transform of the film IR absorption spectrum.

morphology of the film produced at low substrate temperatures (in the process of deposition, the temperature T_s varied from 50 to 150°C), in the case in which the carrier gas was used. The IR absorption spectrum corresponding to this structure is shown in Fig. 1b. The film spectrum differs from that of the initial crystalline PTFE only by the presence of a shoulder at 1000 cm^{-1} and a slight absorption at 745 cm^{-1}. These facts indicate an elevated concentration of CF_3 end groups, i.e., the lower molecular weight of structural units compared to the initial material. Absorption bands of the symmetric (1155 cm^{-1}) and asymmetric (1215 cm^{-1}) extension of CF_2, as well as the effect of bonds in CF_2 groups (641/629 cm^{-1}) correspond to pure crystalline PTFE (linear molecular chains); within the range 1400–4000 cm^{-1}, the absorption is absent [10].

Figure 2a shows the amorphous structure obtained at a high substrate temperature $T_s \approx 290$°C without the use of the carrier gas. The IR absorption spectrum for this structure is presented in Fig. 2b. From a comparison of Figs. 1a, 1b and Figs. 2a, 2b, it follows that the process of structural relaxation at high substrate temperatures is incomplete. The additional absorption bands in Fig. 2b testify to the presence of highly

branched short molecules with a large number of double bonds.

The data obtained in this study are in good agreement with those of [11].

We can emphasize two key features of film deposition from a gas flow by the method described above. First, there is a possibility of using a monomeric initial gas with further polymerization on a substrate and, in part, in the adiabatically expanding supersonic flow. Second, one can realize a unique possibility of controlling the deposition process by varying the parameters that determine the gas flow and the state of a substrate.

ACKNOWLEDGMENTS

We are grateful to Academician V.M. Buznik for his stimulating discussions, to Prof. V.M. Yarygin for his unceasing support of this work, and to R.S. Sharafudinov for his essential help in performing the experiments.

This work was supported by the Presidium of the Russian Academy of Sciences (project no. 8.8), the Council of the President of the Russian Federation for Support of Young Russian Scientists and Leading Scientific Schools (project no. NSh-910.2003.1), and the Russian Foundation for Basic Research (project no. 03-01-00213).

REFERENCES

1. I. I. Bondar', N. P. Dubinin, L. M. Gerdt, *et al.*, Poverkhnost: Fiz., Khim., Mekh., No. 7, 44 (1983).
2. V. P. Kazachenko and A. V. Rogachev, Khim. Vys. Énerg. **34** (4), 270 (1999).
3. A. Pique, R. A. McGill, D. B. Chrisey, *et al.*, Thin Solid Films **353**–**356**, 536 (1999).
4. C. Biloiu, I. A. Biloiu, Y. Sakai, *et al.*, J. Vac. Sci. Technol. A **22** (1), 13 (2004).
5. J. H. Corel, Plast. Technol. **10** (10), 45 (1964).
6. K. K. S. Law, S. K. Murthy, H. G. Pryce Lewis, *et al.*, J. Fluorine Chem. **122**, 93 (2003).
7. S. L. Madorsky, *Thermal Degradation of Organic Polymers* (Interscience Pub., Wiley, 1964; Mir, Moscow, 1967).
8. E. E. Lewis and M. A. Naylor, J. Am. Chem. Soc. **69** (8), 1968 (1947).
9. A. K. Rebrov, J. Vac. Sci. Technol. A **19**, 1679 (2001).
10. N. P. G. Roeges, *A Guide to the Complete Interpretation of Infrared Spectra of Organic Structures* (Wiley, Chichester, 1994).
11. H. Usui, H. Koshikawa, and K. Tanaka, J. Vac. Sci. Technol. A **13**, 2318 (1995).

Translated by Yu. Vishnyakov

Doklady Physics, Vol. 49, No. 12, 2004, pp. 743–746. Translated from Doklady Akademii Nauk, Vol. 399, No. 4, 2004, pp. 486–489.

MECHANICS

Mechanism of Generating Self-Excited Oscillations in Swirling-Jet Effluxes

D. G. Akhmetov and V. V. Nikulin

Presented by Academician V.M. Titov May 18, 2004

Received July 6, 2004

In effluxes of submerged swirling jets into open space, oscillation phenomena often arise. These phenomena manifest themselves as strong periodic pulsations of pressure and velocity, which lead to acoustic radiation (e.g., in air). The results of the first [1] and later [2–6] systematic investigations in this field of research are presented in the literature. However, although the above effects have been known for a long time, the origin of these periodic pulsations, as well as their generation mechanism, have remained poorly understood up to now. Such a state of affairs is explained by the complicated, three-dimensional, and unsteady character of the flows. The most widely known hypotheses on the generation mechanism are the precession of the entire vortex core about the vortex-chamber symmetry axis [4], or the rotation of the vortex core that has been twisted into a spiral after its escape from the chamber [5]. It was also assumed that a counterflow from the open space into the chamber is concentrated near the symmetry axis, where an immobile (rest) point can appear due to the collision of flows coming from inside and outside the chamber [3].

In this study, we have found a new generation mechanism for the oscillations. It is established that, in the case of efflux in the form of a submerged swirling jet through a nozzle from a vortex chamber, the vortex core that coincides with its axis inside the chamber sharply deviates in the nozzle, away from the symmetry axis and towards the nozzle wall. As a result, a bend resembling the end of a hockey stick is formed. The bent part of the vortex core rotates around the symmetry axis of the chamber. This flow structure can be explained by the fact that, unlike in the previous concepts, the rest point in the case of the efflux of the swirling jet into the open space is shifted in the radial direction from the symmetry axis. As a result, both the rest point and the vortex bend rotate around the symmetry axis, producing periodic pulsations of the flow parameters.

Lavrent'ev Institute of Hydrodynamics, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 15, Novosibirsk, 630090 Russia

1. Experimentally, a swirling jet can be produced as a result of fluid outflow from a vortex chamber. Such a formation is an imitation of typical flows that arise in various swirling systems. A vortex chamber is a hollow cylinder, one of whose ends is blocked off while the other terminates in a nozzle. The profile of the nozzle inside the chamber has the shape of a convergent 10-mm-long confuser that further smoothly passes to a 10-mm-long cylindrical hole. The chamber and the nozzle are shown in Figs. 1, 2, and 4. A fluid is delivered into the vortex chamber tangentially to the chamber surface through six identical slit channels made on the cylindrical-chamber surface near the closed end. The entire design has axial symmetry. The diameter of the vortex chamber is 28 mm, and the slit widths are 2 mm. In the basic series of our experiments, the lengths of the chamber and of the slits were 77 and 6.7 mm, respectively. The output diameter of the nozzle was 14 mm. In this design, there is a possibility of replacing the nozzles and varying the lengths of the tangential inputs and of the chamber itself. The experiments were performed in water. The vortex chamber was installed vertically, with the nozzle directed upward. The fluid exited this chamber into a vessel with the shape of a rectangular parallelepiped ($180 \times 180 \times 300$ mm^3 in size), from which, in turn, the fluid exited through drains in the upper part of the vessel. The dimensions of the vessel were sufficiently large so that the fluid in the regions outside the jet was virtually immobile. In our experiments, we studied the unsteady motion of the vortex core and the qualitative structure of the flow near the nozzle output. The mass fluid-flow rate was determined. From this quantity and the input-slit areas, the input flow velocity was calculated. Finally, based on the known flow velocity, the chamber radius, and the kinematic viscosity, we found the Reynolds number.

2. The position of the vortex core was visualized using small air bubbles. The bubbles were introduced into the chamber through either a small hole 0.6 mm in diameter at the closed-end center or via tangential inputs along with the fluid. This method is based on the assumption that, due to the fluid rotation, the pressure near the vortex-core axis is lowered. If the pressure is

1028-3358/04/4912-0743$26.00

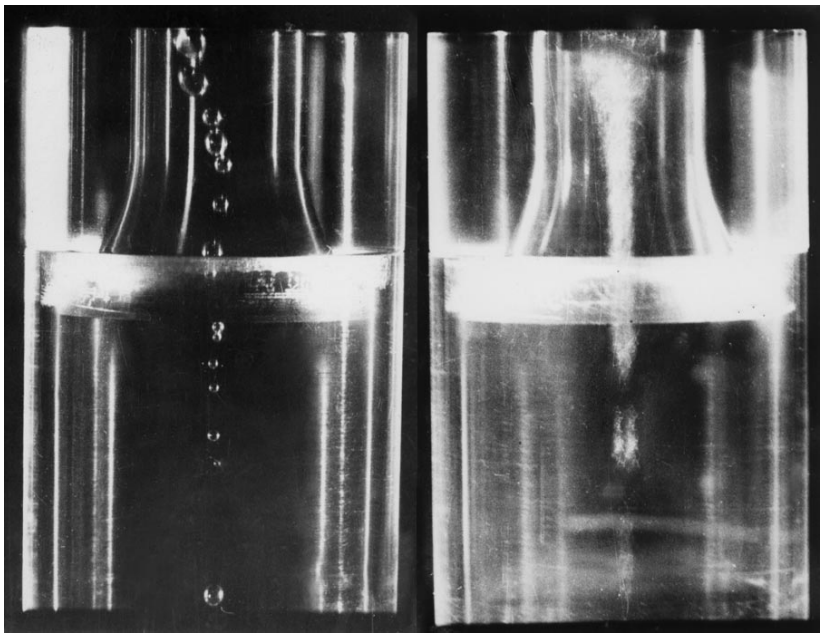


Fig. 1. Air bubbles visualizing the vortex axis. The photographs correspond to different exposures: (left) with the flash lamp and (right) with a long (2 s) exposure time.

low enough, then the bubbles are collected near the vortex axis, thereby visualizing its position.

In the basic series of our experiments, we observed the following phenomena. For $\mathrm{Re} < 6 \times 10^3$, bubbles leave the core rather rapidly. For $\mathrm{Re} \approx 6 \times 10^3$, a single bubble is trapped near the output, where this bubble then rotates around the symmetry axis. A similar pattern was also observed in [1]. The number of trapped bubbles increases with the Reynolds number, and they align in a row. For $\mathrm{Re} > 1.4 \times 10^4$, the bubbles merge, forming a continuous near-axial cavity. Figure 1 shows photographs of bubbles obtained in the same conditions $\mathrm{Re} = 7.5 \times 10^3$: the left and right pictures correspond to exposures made with a flash lamp (exposure time ~ 1–2 ms) and with a long exposure (exposure time ~ 2 s), respectively.

Figure 2 presents a photograph of the vortex that was obtained with the flash lamp upon introducing a colorant through the hole at the center of the closed chamber end. It can be seen from Fig. 1 that the vortex core undergoes a sharp bend inside the nozzle, whereas its shape prior to and beyond the nozzle is virtually a straight line. It is also seen that, immediately beyond the nozzle, the core is absent, since the bubbles rather rapidly leave this region and float upward. Thus, in this region a sharp pressure jump takes place. This jump serves as a barrier to the upward float of the bubbles. The decay of a vortex immediately after the nozzle is also seen in Fig. 2. The right photograph of Fig. 1 also shows that, after curving (core bend), a part of the core rotates around the symmetry axis to form a conical figure visible in the photograph. Before bending, the core is rectilinear and virtually coincides with the rotation axis. Using a stroboscope, we have established that the bend rotates with a constant angular velocity. Under the considered conditions, the rotation frequency was 17 Hz.

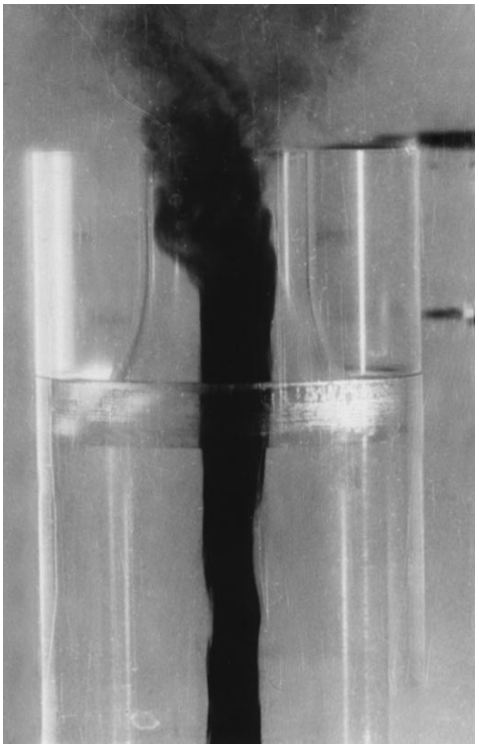


Fig. 2. Vortex photograph colored by an ink delivered at the center of the vortex-chamber bottom boundary.

In this case, core precession was not observed inside the chamber, and the bubbles only slightly drifted along the rotation axis (Fig. 1, right photograph). Using a hydrophone mounted on the nozzle exit section (immediately near the nozzle output), the existence of pressure pulsations was established, and the frequency at which the pulsation amplitude was maximal was determined. This frequency coincides with the core-bend rotation frequency. It was also found that this frequency does not vary either in the presence or in the absence of bubbles. Thus, the bubbles do not noticeably affect the flow. With an increase in the Reynolds number (when a continuous cavity forms along the vortex axis), the pattern remains similar to that shown in Fig. 1. The cavity that is rectilinear inside the chamber is sharply deviated in the nozzle, and the bend rotates with a constant angular velocity.

3. To explain the flow features observed, it is necessary to model its kinematic structure, which qualitatively differs from those proposed previously [3, 4]. In the development of such a model, it was taken into account that the bend rotates with a constant angular velocity, and that the collision of the jet flowing out of the chamber with the counterflow from the open space occurs. The existence of the counterflow was known earlier [3] and was confirmed in our experiments described above. As a result, the instantaneous pattern of streamlines before the vortex decay in the reference system lying in the vortex-axis plane and rotating along with the bend can be qualitatively presented in schematic form in Fig. 3. The symmetry axis is shown by the *OO* line and the rest point is *a*. Note that this structure is not completely immobile but executes small oscillations in both the axial and radial directions, which is confirmed by observations of the bubbles. In addition, in the laboratory reference system, this pattern rotates around the *OO* axis. As follows from Fig. 3, the vortex core and the larger part of the fluid flowing out of the chamber along its axis are deviated in one direction.

To verify the flow structure presented here, we performed the following experiment. A thin tube 0.6 mm in diameter was placed along the continuation of the chamber axis at a distance of 4 mm from the nozzle exit section. Through this tube, a colorant was slowly introduced. Figure 4 shows a photograph showing the colorant outflow from the tube. The photograph was obtained with illumination by a flash lamp. It is seen that the colored line is deviated from the chamber axis in one direction. In the laboratory reference system, the line rotates around this axis and qualitatively corresponds to the streamlike *bc* in Fig. 3. Note that, since the precession of the vortex core before the appearance of the bend was not observed in our experiments, this implies that the vortex-core axis prior to the bend was close to or coincided with the *OO* line. Otherwise, precession would be observed. Experiments performed with nozzles 8 and 20 mm in diameter yield results qualitatively similar to those presented above.

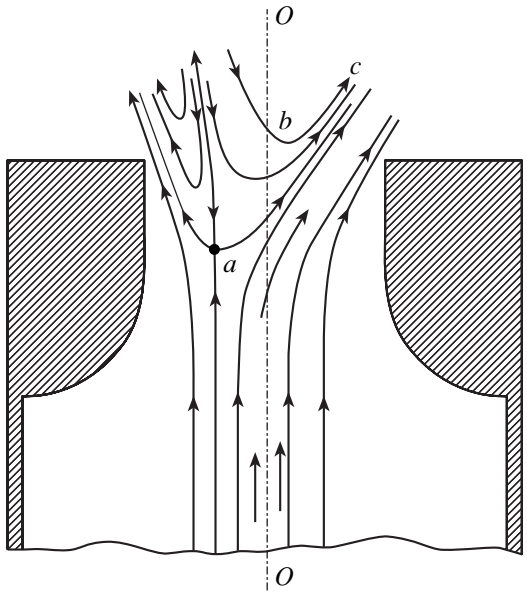


Fig. 3. Schematic pattern of streamlines in the plane of the vortex axis in the rotating frame of reference.

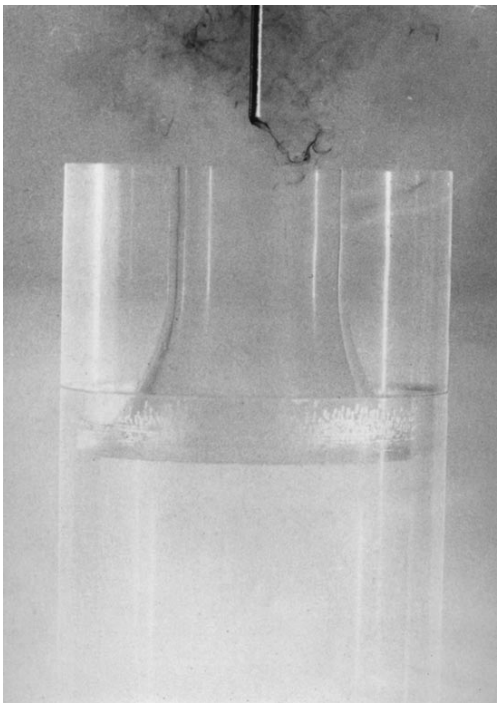


Fig. 4. Flow at the nozzle exit section. The visualization is made by an ink delivered to the symmetry axis. The photograph was taken with the illumination by a flash lamp.

Thus, in this study, we have discovered the possibility of the existence of flows with a sharp bending of the linear vortex that rotates at a constant angular velocity. In order to explain the phenomenon observed, we have proposed a flow structure in which the rest point, in the

case of the collision of a swirling jet with the counterflow, is not located on the symmetry axis: it is shifted from it in the radial direction. As a result, the rest point and the bend rotate around the axis. Note that the symmetry violation observed takes place even in the presence of axial symmetry of the chamber; i.e., it is a property of the rotating flow itself. The results obtained provide qualitatively new insights into strongly swirling flows.

ACKNOWLEDGMENTS

This work was supported in part by Integration project no. 25, Siberian Division, Russian Academy of Sciences.

REFERENCES

1. B. A. Vonnegut, J. Acoust. Soc. Am. **26**, 18 (1954).
2. R. C. Chanaud, J. Acoust. Soc. Am. **35**, 953 (1963).
3. R. C. Chanaud, J. Fluid Mech. **21**, 111 (1965).
4. Yu. A. Knysh and S. V. Lukachev, Akust. Zh. **23**, 776 (1977).
5. A. Sh. Kiyasbeĭli and M. E. Perel'steĭn, *Vortex Fluid Flowmeters* (Mashinostroenie, Moscow, 1974) [in Russian].
6. D. G. Akhmetov and V. V. Nikulin, Izv. Akad. Nauk, Ser. Mekh. Zhidk. Gaza, No. 3, 60 (2004).

Translated by Yu. Vishnyakov

Doklady Physics, Vol. 49, No. 12, 2004, pp. 747–750. Translated from Doklady Akademii Nauk, Vol. 399, No. 4, 2004, pp. 490–493.

MECHANICS

Experimental Observation of the Solutocapillary Migration of Air Bubbles in Inhomogeneous Liquid Solutions

A. L. Zuev and K. G. Kostarev

Presented by Academician V.P. Matveenko April 11, 2004

Received April 14, 2004

INTRODUCTION

The ability of air bubbles in a liquid to move spontaneously toward a decrease in the surface tension [1] can have a significant effect on various technological processes occurring in multicomponent media. This is especially important under conditions when the gravitational mechanisms responsible for motion are weak or absent (for example, in microgravity) [2]. Usually, this motion arises in nonuniformly heated liquids due to the temperature dependence of the surface tension coefficient σ. In this case, the effect is called thermocapillary migration of bubbles [3]. Meanwhile, in nonuniform-composition media (in particular, in binary solutions of liquids), the other reason for changing the surface tension σ can exist. We imply the dependence of surface tension on the concentration of a dissolved substance. In such a situation, the solutocapillary migration, i.e., motion of bubbles toward the more surfactant of a solution, will be observed. Experimental examination of the phenomenon is much hampered by both subsidiary effects associated with the presence of gravity and difficulties in producing a stationary concentration gradient. It should also be noted that the methods of measuring such gradients on the free surface still remain poorly developed. Therefore, the solutocapillary migration of bubbles, which is predicted theoretically, was not previously observed in experiments.

In this study, under ground-based laboratory conditions, we have discovered and examined the migration of air bubbles in inhomogeneous aqueous methanol solutions. With this purpose in mind, we have developed a number of methods making it possible to suppress the action of gravitational mechanisms of motion and produce quasisteady longitudinal concentration gradients of a surfactant in thin horizontal layers of liquid. We have measured the dependences of the migration velocity on the bubble size, time, concentration gradient, and parameters of liquid. It was established that, unlike the thermocapillary effect, the solutocapillary migration rather rapidly decays with time. We assumed that this can be caused by the adsorption of a surfactant on the free surface of a bubble.

EXPERIMENTAL METHOD

Under laboratory conditions, the motion of bubbles is mainly hampered by the buoyancy force, which causes bubbles to ascend in a vertical direction. Therefore, to neutralize the effect of gravity, we used a technique similar to that applied by us previously in studies of thermocapillary bubble migration in nonuniformly heated liquid [4]. Small air bubbles were introduced into a thin horizontal layer of liquid having a longitudinal concentration gradient. The layer was bounded at the top and bottom by solid surfaces. In this case, the bubbles turned out to be pressed by the buoyancy force to the upper boundary of the layer. Therefore, they were able to move only horizontally toward the concentration gradient. Small bubbles with a diameter less than the layer thickness retained their spherical shape, whereas bubbles of a larger size take the shape of a cylindrical pellet flattened out between the horizontal walls of a cell.

Of course, in such a formulation of the problem, the interaction between bubbles and the solid boundaries of the liquid begins to play a significant role. The necessary condition of bubble motion is the presence of a thin interlayer between the bubble ends and the cell walls. This can be attained only in the case in which the solid surface is completely wetted by the liquid (wetting angle is close to zero). Then, a bubble does not adhere to the walls and acquires the ability to freely slide along them. The existence of such a thin film was confirmed by special interferometric experiments. Another problem in the experiments is the impossibility of maintaining a given concentration gradient. This is stipulated by the fact that the maintenance of a concentration difference assumes the introduction and, consequently, removal of the mixture at the opposite ends of a cell, which causes undesirable additional flows. We man-

Institute of Mechanics of Continua, Ural Division, Russian Academy of Sciences, ul. Akademika Koroleva 1, Perm, 614013 Russia

e-mail: vab@icmm.ru

1028-3358/04/4912-0747$26.00

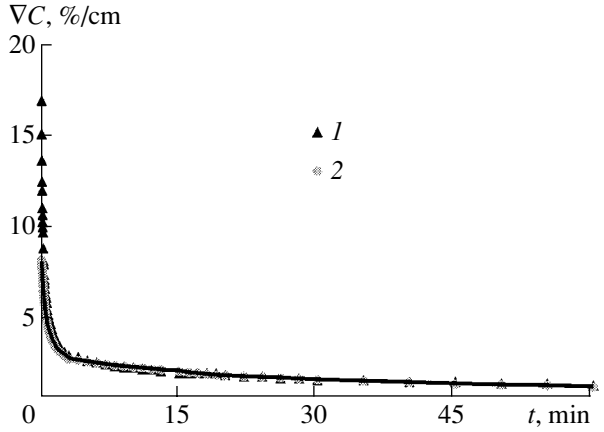


Fig. 1. Time dependence of the concentration gradient at the cell center.

aged to produce only a quasi-steady concentration gradient, the value of which decreases with time due to diffusion. Hence, the necessity of permanently determining the concentration fields of surfactant by optical methods follows.

The cell used in our investigations was an interferometric cell made in the form of rectangular cavity with sizes of $90 \times 40 \times 1.2$ mm. Its walls were made of plane-parallel glass plates with semitransparent mirror coating. The observation was performed in reflected light on the wide-facet side. A Fizeau laser interferometer made it possible to visualize the concentration distribution within the liquid in the form of isolines of the refractive index whose variation is proportional (in the isothermal case) to the concentration change [5]. The cell was filled with aqueous solutions of methanol with concentrations of C_0 between 80 and 100% (by concentration, we imply the mass fraction of a surfactant dissolved in water). Methanol solutions were chosen after test experiments with various liquids. Preliminary experimental studies have shown that, in the majority of liquids, the arisen solutocapillary flow turns out to be too weak in order to overcome the friction forces and cause bubble motion. At the same time, in experiments with methanol solutions, the solutocapillary migration of bubbles was clearly seen, owing to a very fortunate combination of physicochemical methanol properties. Methanol satisfies the main requirements specific for working liquids: total wetting, transparency, nonaggression properties, and density close to that of water. In addition, it is characterized by the least viscosity and surface tension [6]. These properties characterize it as an extremely strong surfactant with respect to water. In combination with low viscosity, this considerably intensifies the solutocapillary flows. Within the range of chosen concentrations, the methanol solutions have nearly linear concentration dependences of density, viscosity, surface tension, and refractive index. All the experiments were performed in the isothermal regime at a temperature of (20 ± 1)°C.

The longitudinal concentration gradient in the horizontal layer of the solution was produced according to the following method [7–9]. First, the experimental cell was installed vertically on a small facet and was half filled with the solution. Thereafter, the lighter pure alcohol was added from above. 10–20 s later, a distribution of the concentration stratified in density was established in the cell (due to the gravitational force). The distribution represented two areas of initial liquids positioned under each other and separated by a narrow zone of diffusion transition between them. Because of the smallness of the diffusion coefficient, such a gravitationally stable vertical concentration distribution was able to exist for several hours. As soon as the convective motion associated with the process of pouring the liquids into the cell was completed, it was turned into the horizontal position. As a result, there appeared a considerable longitudinal density difference leading the system of the liquids to a convective shear motion. The intensity of this flow rapidly decreased with time by virtue of the small layer thickness and active mixing of alcohol and its solution in counter flows. As a result, at the cell center, an extended area of low but rather homogeneous longitudinal concentration gradient was formed. The area was bounded by narrow zones of the initial liquids near the cell ends. The transition from one monotonous band to the other in the interference pattern corresponded to the variation in the alcohol concentration by about 0.3%, which allowed us to calculate the value of the concentration gradient.

Figure 1 shows the time dependences of the longitudinal concentration gradient at the center of the cell in its horizontal position. In the cases (*1*) and (*2*), the layer of pure methanol was poured above the methanol solution with the concentration 80% and 90%, respectively. As is seen, the curves behave similarly and virtually merge into a single curve despite of the 10% difference in the initial concentrations. At first, the gradient very rapidly dropped to a certain value (approximately, to 2.5%/cm) due to convective motion. However, further, it slowly decreased for a long period of time (tens of minutes), which corresponded to a quasi-diffusion regime of dissolution. It is this nearly linear quasisteady horizontal concentration gradient that was used by us to study the solutocapillary bubble migration. Note that, in experiments with cells of greater thicknesses (~2 mm and thicker), the convective-flow process was completely different and the longitudinal concentration gradient was not attained. Upon turning such cells (after they had been filled with the liquids) from the vertical position to a horizontal one, the leaking of the lighter layer onto the heavier layer was observed, and there appeared a two-layer system of the liquids with vertical concentration stratification. Thus, the small thickness (~1 mm) of the liquid layer is proven to be one of the main requirements for successfully performing the experiments.

RESULTS

Using a medical syringe, we introduced an air bubble into the layer of the solution in which the horizontal concentration gradient was produced, therewith, the bubble took the shape of a plane disk 5–15 mm in diameter. As soon as the bubble was separated from the needle, it began to move toward the higher alcohol concentration. Simultaneous visual observation of the concentration field and the bubble position in it made it possible to determine the concentration gradient, diameter, and bubble-migration velocity at different instants of time. A typical interference pattern is presented in Fig. 2. The experiments performed showed that, as in the case of thermocapillary migration [4], the maximum velocity of the solutocapillary motion is proportional to both the surface-tension gradient and bubble diameter. The results of experiments with bubbles of various sizes for different values of the concentration gradient are shown in the summary plot (Fig. 3), where the ratio of maximum migration velocity u to bubble diameter D is plotted along the ordinate axis.

At the same time, the experiments have demonstrated that the solutocapillary and thermocapillary mechanisms of the bubble migration significantly differ. For example, in the uniform temperature gradient, a bubble moved with a constant time-independent velocity. However, in the concentration field, the bubble velocity monotonously dropped to zero for 40–60 s, in spite of the fact that the concentration gradient in the solution surrounding the bubble remained rather high. However, a new bubble inserted into the solution near the bubble, which was at rest, began to move. A typical time dependence of the migration velocity is shown in Fig. 4 (concentration gradient is 2.5%/cm and D = 7 mm). With varying the bubble size, the situation remained qualitatively unchanged, although the migration time and path length increased with increasing the bubble diameter.

To explain a decrease in the migration velocity, we can assume that, in the process of bubble motion, the bubble surface is saturated by alcohol due to adsorption from the solution. The bulk diffusion of the surfactant fails to entirely counteract the adsorption in producing a uniform concentration distribution on the free surface. As a result, the surface tension on the entire boundary of the bubble rather rapidly equalizes, regardless of the concentration gradient in the surrounding liquid. Therefore, the cause responsible for the bubble migration disappears. Hence, it follows that, in other cases in the presence of a surfactant dissolved in liquid, the action of the solutocapillary effects can be observed only for a limited time determined by the rate of the adsorption process.

In our opinion, the adsorption phenomenon is capable of explaining the well-known fact that, in liquids containing impurities of surfactant (in particular, in high-concentrated aqueous solutions of organic liquids), the thermocapillary convection is usually not observed [10]. For example, in our experiments [9] in which we studied the thermocapillary migration in nonuniformly heated aqueous solutions of methanol, even after adding to the alcohol an insignificant amount of water (i.e., in the 90% solution that only insignificantly

Fig. 2. Interferogram of the concentration field in the process of the bubble migration.

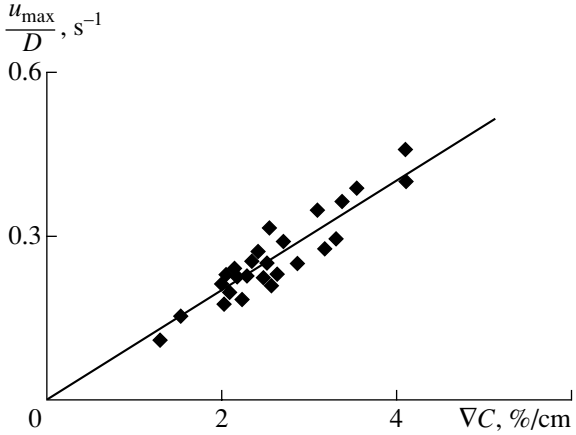


Fig. 3. The ratio of the solutocapillary migration velocity normalized to the bubble diameter as a function of the methanol concentration gradient.

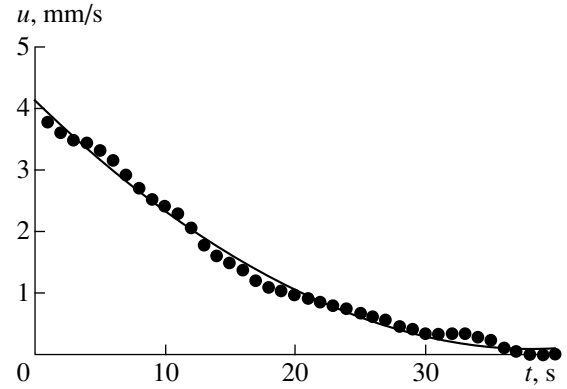


Fig. 4. Time dependence of bubble-motion velocity.

differs from the pure alcohol in its physicochemical properties), the bubble migration was not observed. Most likely, the reason is that the thermal nonuniformity of the surface tension is compensated by the redistribution of surfactant on the bubble surface. Since the typical diffusion times considerably exceed the thermal times, the reasons responsible for motion of liquid, namely, the nonuniformity of the surface tension, disappear, as is observed in the purely concentration case.

ACKNOWLEDGMENTS

We are grateful to Yu.K. Bratukhin and A.F. Pshenichnikov for fruitful discussions. This work was supported by the Russian Foundation for Basic Research, project no. 03-01-00579.

REFERENCES

1. N. O. Young, J. S. Goldstein, and M. J. Block, J. Fluid Mech. **6**, 350 (1959).
2. *Hydromechanics of Weightlessness*, Ed. by A. D. Myshkis (Nauka, Moscow, 1976).
3. Yu. K. Bratukhin, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 5, 156 (1975).
4. Yu. K. Bratukhin and A. L. Zuev, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 3, 62 (1984).
5. B. I. Burshteĭn, A. L. Zuev, Z. P. Podkovyrkina, *et al.*, in *Numerical and Experimental Modeling of Hydrodynamical Phenomena in Weightlessness* (Ural. Otd. Akad. Nauk SSSR, Sverdlovsk, 1988), p. 108.
6. *Handbook of a Chemist*, Ed. by B. P. Nikol'skiĭ (Khimiya, Leningrad, 1962), Vol. 3.
7. A. L. Zuev and A. V. Shmyrov, in *Annotation Reports of 8th All-Russia Conference on Theoretical and Applied Mechanics* (Perm, 2001), p. 282.
8. A. L. Zuev, V. A. Briskman, K. G. Kostarev, *et al.*, in *Abstracts of I International Marangoni Conference* (Giessen, 2001), p. 4.
9. A. L. Zuev, K. G. Kostarev, and A. V. Scmyrov, in *Thermocapillary and Solutocapillary Effects in Complex Systems* (Ural. Otd. Akad. Nauk, Yekaterinburg, 2003), p. 105.
10. V. G. Levich, *Physicochemical Hydrodynamics* (Fizmatgiz, Moscow, 1959).

Translated by Yu. Vishnyakov

Doklady Physics, Vol. 49, No. 12, 2004, pp. 751–755. Translated from Doklady Akademii Nauk, Vol. 399, No. 5, 2004, pp. 620–624.

MECHANICS

Multimode Bifurcation of the Flow of a Viscous Fluid in a Plane Diffuser

L. D. Akulenko and S. A. Kumakshev

Presented by Academician V.F. Zhuravlev May 31, 2004

Received June 1, 2004

The complete pattern of steady flows of a viscous incompressible fluid in a plane diffuser with a fixed expansion angle has been determined and analyzed (the Jeffery–Hamel problem). For Re < 0, we show that there is a countable set of finite adjoining intervals where the complex bifurcation of flows occurs. When Re decreases, the number of modes increases in flows containing inflow and outflow domains of the fluid. It is shown that the velocity profiles of odd- and even-mode flows are symmetric and asymmetric about the diffuser axis, respectively. With the decrease $\mathrm{Re} \longrightarrow -\infty$ and the corresponding increase in the interval number, the interval length and the number of allowed flows increase unboundedly. These results clarify the mechanism of the phenomenon under consideration—the loss of the stability of laminar diffuser flows at the initial stage of turbulization.

1. We consider the steady flow of a viscous incompressible fluid in a plane diffuser

$$\Omega = \{(r, \theta): r > 0, |\theta| < \beta\}, \quad 0 < \beta \le \pi, \qquad (1)$$

where Ω is the flow domain, (r, θ) are the polar coordinates, and β is the half-expansion angle. On the diffuser walls, $r > 0$ and $\theta = \pm\beta$, no-slip conditions are valid, and the flow has a source-type singularity at $r = 0$. For the classical Jeffery–Hamel problem [1–3], the source strength $Q < 0$ is given; i.e., the fluid rate through any section $r > 0$ in the domain Ω given by Eq. (1) is fixed.

Only two dimensionless parameters—the expansion angle 2β and the Reynolds-number analog $\mathrm{Re} = \dfrac{\rho Q}{\mu} < 0$, where ρ is the volume density and μ is the dynamic viscosity of the fluid—can be introduced for the system. These are insufficient for making the equations of motion completely dimensionless [3–5]. However, one can construct a self-similar solution with the radial velocity field

$$v_r = -\frac{Q}{r}V(\theta), \quad v_\theta \equiv 0, \quad (r, \theta) \in \Omega, \qquad (2)$$

where V is an unknown dimensionless function. This solution satisfies the incompressibility condition for the arbitrary smooth function $V(\theta)$.

The Navier–Stokes equations, no-slip boundary conditions, and the constant-rate condition provide the constitutive relationships [1–3]

$$V'' + 4V - \mathrm{Re}V^2 = C = \mathrm{const}, \quad |\theta| < \beta;$$
$$V(\pm\beta) = 0, \quad \int_{-\beta}^{\beta} V(\theta)d\theta = 1. \qquad (3)$$

Boundary and integral conditions (3) determine the integration constants and the constant C. The pressure corresponding to the solution given by Eq. (2) is expressed in terms of V and C as

$$p = \frac{\rho Q^2}{2r^2\mathrm{Re}}(C - 4V), \quad (r, \theta) \in \Omega. \qquad (4)$$

The nonlinear boundary value problem given by Eq. (3) with the additional integral condition is symmetric under the $\theta \longrightarrow -\theta$ transformation. However, in addition to symmetric flows [1–7], this problem has multimode solutions asymmetric about the confuser axis, as we showed for confuser (Re > 0) [8] and diffuser (Re < 0) [9] flows. By introducing the unknown Z, which characterizes the fluid rate, and differentiating the equation for V, the boundary value problem with integral condition (3) is reduced to the equivalent nonlinear fourth-order boundary value problem

$$V''' + 4V' - 2\mathrm{Re}VV' = 0,$$
$$Z' = V - \frac{1}{2\beta}; \quad V(\pm\beta) = Z(\pm\beta) = 0. \qquad (5)$$

A solution of problem (5), i.e., V, V', V'', and Z, as well as $\mathbf{v}(r, \theta)$ and other kinematic and dynamic char-

Institute for Problems in Mechanics,
Russian Academy of Sciences,
pr. Vernadskogo 101-1, Moscow, 119526 Russia
e-mail: kumak@ipmnet.ru

1028-3358/04/4912-0751$26.00

acteristics of flows, must be found as functions of r, θ, and the determining parameters β and Re for the entire domain Ω given by Eq. (1) for $\mathrm{Re} < 0$ and $0 < \beta \le \pi$.

2. The formal analytical solution of problem (3) by means of the first integral leads to elliptic functions and integrals [2, 3, 6, 7]. A system of two transcendental equations containing two parameters (β and Re) can be obtained for determining unknown constants. This solution is implicit, and its use for particular calculations leads to substantial calculation difficulties. These difficulties are caused by the degenerate nature of equations for $\mathrm{Re} \longrightarrow \pm 0$ and $\mathrm{Re} \longrightarrow \pm\infty$, as well as for a countable set of negative Re values. Moreover, the solution as a function of the angle β is also irregular [5]. Calculations are further complicated because the reference data for elliptic functions and integrals are insufficiently accurate. The available results are obtained by solving the boundary value problem given by Eq. (3) without the integral condition (the rate of the inflow or outflow of the fluid on the diffuser or confuser axis is given).

To solve the nonlinear boundary value problems given by Eqs. (3) and (5), we developed a constructive numerical–analytical method [10]. It is based on a fast-convergence algorithm and continuation in the parameters. For convenience, the normalized unknowns y and z, the argument x, and the parameters γ and λ, which are given by

$$y(x) = 2\beta V(\theta), \quad z = 2\beta Z, \quad \lambda = 8\beta^3 C,$$
$$x = \frac{1}{2}\left(\frac{\theta}{\beta} + 1\right), \quad 0 \le x \le 1, \quad \gamma = y'(0) \tag{6}$$

are introduced into Eq. (3). The unknown functions and parameters are determined by solving the following nonlinear boundary value problem for the fixed essential parameters $a = 4\beta$ and $b = 2\beta\mathrm{Re}$:

$$y'' + a^2 y - by^2 = \lambda, \quad y(0) = 0, \quad y'(0) = \gamma;$$
$$z' = y - 1, \quad z(0) = 0; \tag{7}$$
$$y(1) = z(1) = 0.$$

Let the solution of the boundary value problem given by Eq. (7) be known for $a = a_0$ and $b = b_0$. In this case, the fast-convergence algorithm [10] is applicable for $a = a_0$ and $b = b_0 + \delta b$, where δb is sufficiently small. In this method, the unknowns γ and λ are refined by the recursive procedure

$$\gamma_{n+1} = \gamma_n + \delta\gamma_n, \quad \lambda_{n+1} = \lambda_n + \delta\lambda_n,$$
$$\delta\gamma_n = -[y_n(1)s_n(1) - z_n(1)h_n(1)]\Delta_n^{-1}(1),$$
$$\delta\lambda_n = [y_n(1)w_n(1) - z_n(1)g_n(1)]\Delta_n^{-1}(1), \tag{8}$$
$$\Delta_n(x) = y_n(x)s_n(x) - h_n(x)w_n(x),$$
$$\Delta_n(1) \ne 0, \quad n = 0, 1, \ldots.$$

The functions $y_n(x)$ and $z_n(x)$ are solutions of the Cauchy problem specified by Eq. (7) for given $\gamma = \gamma_n$ and $\lambda = \lambda_n$, and g_n (w_n) and h_n (s_n) are the sensitivity functions of the solution y_n (z_n) to the parameters γ and λ, respectively. They are obtained by high-accuracy integration of the linear Cauchy problems

$$g'' + a^2 g - 2byg = 0, \quad w' = g;$$
$$g(0) = w(0) = 0, \quad g'(0) = 1;$$
$$h'' + a^2 h - 2byh = 1, \quad s' = h; \tag{9}$$
$$h(0) = h'(0) = s(0) = 0,$$

where $a = a_0$, $b = b_0 + \delta b$, and $y = y_n(x)$.

Calculations are terminated for $n = n^*$ when the required absolute and/or relative accuracy, which is determined by the residuals in $y_n(1)$ and $z_n(1)$, is reached. Then, the parameter b is increased and the procedure specified by Eqs. (7)–(9) is repeated. Thus, the curves $\gamma(a_0, b)$ and $\lambda(a_0, b)$ can be obtained, and their properties can be analyzed in a wide range of the parameter b ($b < 0$, i.e., $\mathrm{Re} < 0$).

A similar successive-refinement procedure is applicable to the solution for other values of the parameter $a = a_0 + \delta a$. This procedure allows for simultaneous continuation in the parameters a and b. Along with the residuals in the solutions $y_n(1)$ and $z_n(1)$, residuals in the abscissas $\varepsilon_n = 1 - \xi_n$ and $\mu_n = 1 - \eta_n$, where ξ_n and η_n are the zeros of the functions $y_n(x)$ and $z_n(x)$, respectively, that are closest to the point $x = 1$, may be used to estimate the error and to refine the solutions [10]. Calculations show the fast convergence and efficiency of the algorithm, which provides for accurate operative mass calculations (see below). These calculations for various β (i.e., a) and $b = 2\beta\mathrm{Re}$ values provide the functions $\gamma(b)$ and $\lambda(b)$. Using these functions, one can determine the velocity profile $y(x)$ as a solution of the Cauchy problem given by Eqs. (7), calculate all the kinematic and dynamic characteristics of the flows, and perform a complete qualitative analysis of flows in the diffuser, which is similar to the previous analysis for the confuser [5, 8].

3. The complete pattern of the dependence of the parameters γ and λ on $0 < \beta \le \pi$ and $b \le 0$ is diverse and difficult to describe in detail. Similarly to the case for confuser flows [5], it is convenient to divide the total β range into several intervals corresponding to different behaviors of $\gamma(b)$ and $\lambda(b)$ in the procedure of continuation in the parameter $b < 0$ ($|b| \ll 1$):

$$(1)\ 0 < \beta \le \frac{\pi}{4}, \quad (2)\ \frac{\pi}{4} < \beta \le \frac{\pi}{2}, \quad (3)\ \frac{\pi}{2} < \beta < \beta^*,$$
$$(4)\ \beta^* < \beta \le \frac{3\pi}{4}, \quad (5)\ \frac{3\pi}{4} < \beta \le \pi; \tag{10}$$
$$\beta^* = \frac{1}{2}\arg(\tan\varphi - \varphi) \approx 2.2467047.$$

The solution for $b = 0$ (Re = 0) does not exist for the critical value β^*, and the solution for $b \neq 0$ can be continued in the parameter β for $\beta \lessgtr \beta^*$. The limiting values $\beta \longrightarrow +0$, $\beta^* \pm 0$, and $\beta = \frac{\pi}{2}, \pi$ are interesting for hydrodynamics.

We found that many confuser [5] and diffuser [9] multimode flows that are noncontinuable for Re $\longrightarrow$ 0, along with solutions that are continuable in b, i.e., Re, exist in the β range given by Eq. (10).

It is convenient to illustrate the previously unknown pattern of the bifurcations of steady diffuser flows for Re $\longrightarrow -\infty$ for the case of slightly diverging walls, which is of interest for hydrodynamics and various technologies (usually $\beta = 5°$–$30°$).

Figure 1 shows $\gamma(b)$ and $\lambda(b)$ as calculated by Eqs. (7)–(9) for $\beta = 10° \approx 0.1745$ rad. The behavior of the curves $\gamma(b)$ and $\lambda(b)$ for $0 < \beta < 10°$ is close to the case of $\beta = 10°$. As is seen in Fig. 1, there is a countable set of branches $\gamma^{(n)}(b)$ and $\lambda^{(n)}(b)$ for $n = 1, 2, \ldots, N, N + 1, \ldots$. To describe these branches, it is appropriate to divide the entire range of the parameter $b < 0$ into intervals where flows are qualitatively rearranged with decreasing b, $b \longrightarrow -\infty$.

The first domain of (b, γ) and (b, λ) was analyzed in detail in [9] by using the known analytical solution for $b = 0$ and continuing in b in the range $0 \le b \le b^{(1)*} \approx -21.7$. It has been shown that there are single-mode ($\gamma_1^{(1)}(b)$ and $\lambda_1^{(1)}(b)$) and three-mode ($\gamma_3^{(1)}(b)$ and $\lambda_3^{(1)}(b)$) branches for which flows are symmetric about the diffuser axis. These flows are pure radially diverging flows [6] for the single-mode branch and diverging in the middle (on the axis) and converging (outflow) at the edges of the diffuser for the three-mode branch. Below, the subscript shows the number of modes, i.e., inflow and outflow domains. The mode number of the flow is determined by the number of crossings of the abscissa axis by the corresponding plot $y(x)$ (this is the fluid velocity profile at an arbitrary arc as a function of the angle within the diffuser expansion angle). The integral of the function $y(x)$ is equal to the normalized source strength $Q = 1$. The superscript is the interval number for the parameter $b = 2\beta\text{Re}$. The single- and three-mode regimes exist for $b_*^{(1)} \approx -18.8 \le b \le 0$ and $b_1^{(1)*} \le b \le 0$, respectively. In the range $b_1^{(1)*} < b < b_*^{(1)}$, the curves $\gamma_3^{(1)}$ and $\lambda_3^{(1)}$ are double-valued, and two three-mode flows are therefore possible. One three-mode flow corresponding to $\gamma_3^{(1)}$ and $\lambda_3^{(1)}$ and, as was mentioned above, one single-mode flow exist in the range $b_*^{(1)} \le b < 0$. In addition, two previously unknown two-mode asymmetric flows with the inflow near one wall and the outflow near the other wall exist in this range.

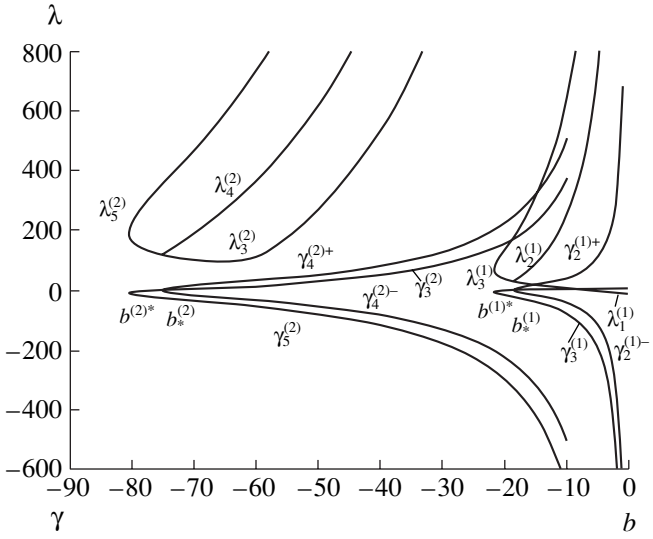


Fig. 1. Plots $\gamma(b)$ and $\lambda(b)$, as well as of the bifurcation pattern for steady flows with decreasing Re. The first and second intervals are $0 < b \le b^{(1)*}$ and $b^{(1)*} < b \le b^{(2)*}$, respectively.

We emphasize that $|\gamma_2^{(1)}|$ and $|\lambda_2^{(1)}|$, as well as $|\gamma_3^{(1)}|$ and $|\lambda_3^{(1)}|$, increase unboundedly when $b \longrightarrow -0$; i.e., these curves cannot be obtained by continuing in the parameter. In particular, $\gamma_3^{(1)} \sim -10^3$ and $\lambda_3^{(1)} \sim 10^4$ for $b = -1$. Thus, the first interval is characterized by two critical b values. The value $b_*^{(1)}$ is a triple point, where the bifurcation of single-, two-, and three-mode flows occurs. In addition, $\gamma_1^{(1)} = 0$ at this point. The value $b^{(1)*}$ is the turning (return) point of the curves $\gamma_3^{(1)}(b)$ and $\lambda_3^{(1)}(b)$ and determines the boundaries of the first interval.

The numerical–analytical procedure specified by Eqs. (7)–(9) reveals the second interval $b^{(2)*} \approx -80.44 \le b \le b^{(1)*}$, where three-, four-, and five-mode flows are possible. The behavior of the curves $\gamma_{3,4,5}^{(2)}$ and $\lambda_{3,4,5}^{(2)}$ is qualitatively similar to the behavior of the curves $\gamma_{1,2,3}^{(1)}$ and $\lambda_{1,2,3}^{(1)}$, respectively. However, in contrast to $\gamma_1^{(1)}$ and $\lambda_1^{(1)}$, which have finite values for $b = -0$, $\gamma_3^{(2)} \longrightarrow -\infty$ and $\lambda_3^{(2)} \longrightarrow \infty$ for $b \longrightarrow -0$. For $b \longrightarrow -\infty$, the curves $|\gamma^{(2)}|$ and $\lambda^{(2)}$ increase much faster. In particular, for $b = -10$, $\gamma_3^{(1)} \approx -70$ and $\lambda_3^{(1)} \approx 10^3$, whereas $\gamma_5^{(2)} \approx -670$ and $\lambda_5^{(2)} \approx 10^4$ (Fig. 1).

The second interval ($n = 2$) has two similar critical points: $b^{(2)*}$ is a turning (return) point for the curves $\gamma_5^{(2)}$ and $\lambda_5^{(2)}$ and $b_*^{(2)}$ is a bifurcation (triple) point near

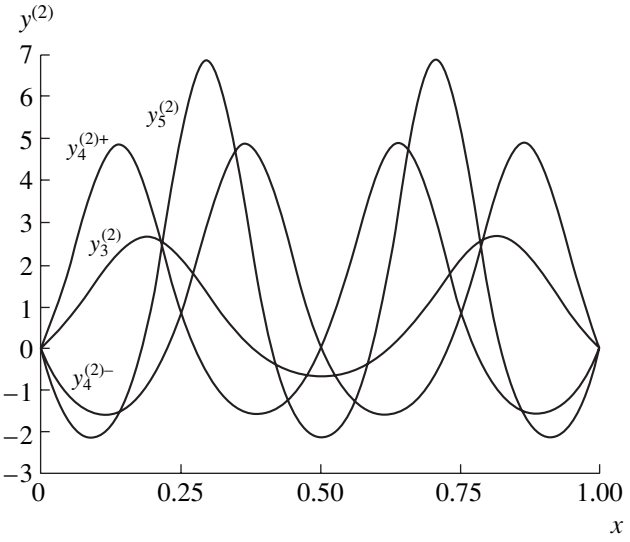


Fig. 2. Velocity profiles $y_{3,5}^{(2)}(x)$ and $y_4^{(2)\pm}(x)$ for $b = -60$ for the second interval.

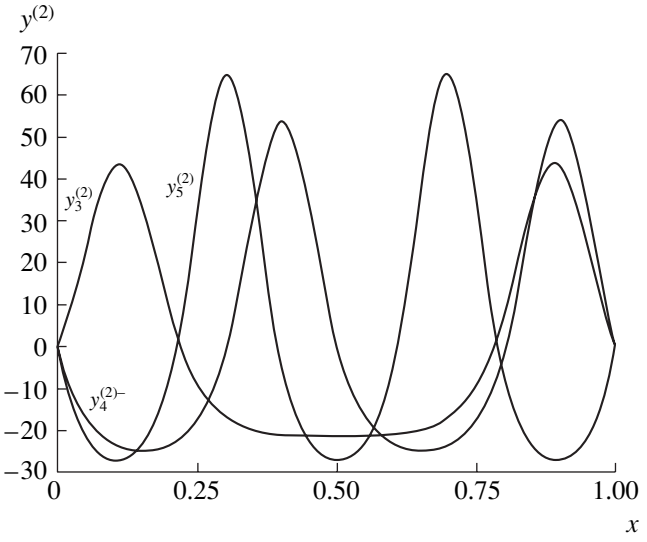


Fig. 3. Velocity profiles $y_{3,5}^{(2)}(x)$ and $y_4^{(2)-}(x)$ ($y_4^{(2)+}$ is omitted) for $b = -10$.

which the three mentioned branches of steady flows exist. In addition, $\gamma_3^{(2)} = 0$ at this point. In the range $b^{(2)*} < b < b_*^{(2)} \approx -75.3982$, the curves $\gamma_5^{(2)}$ and $\lambda_5^{(2)}$ are double-valued, and two five-mode flows correspond to each b value. In the range $b_*^{(2)} \le b < 0$, the curves are single-valued; i.e., one flow of the above form is realized. Four-mode flows (two for $b > b_*^{(2)}$) have alternating inflow and outflow domains. If the inflow exists near one wall, the outflow exists near the other wall. A node, i.e., a zero y value, is present at the middle of the diffuser $\left(x = \frac{1}{2}\right)$. Half the fluid rate passes through each half of the diffuser.

According to Fig. 1, one three-mode flow corresponding to the curves $\gamma_3^{(2)}$ and $\lambda_3^{(2)}$ exists in the range $b_*^{(2)} < b < b^{(1)*}$, two three-mode flows (flow corresponding to $\gamma_3^{(1)}$ and $\lambda_3^{(1)}$ is added) are present for $b = b^{(1)*}$, three three-mode flows are realized in the range $b^{(1)*} < b < b_*^{(1)}$ (because $\gamma_3^{(1)}$ and $\lambda_3^{(1)}$ are double-valued), and two three-mode flows exist in the range $b_*^{(1)} < b < 0$. Other single-, two-, and three-mode flows are not realized in the range $b_*^{(2)} < b < 0$.

The velocity profiles of the flows $y_{1,2,3}^{(1)}(x)$ for the first b interval were obtained and analyzed in detail in [9]. They correspond to the above analysis. Figure 2 shows the functions $y_{3,4,5}^{(2)}(x)$ for $b = -60$ belonging to the second interval. Both four-mode profiles $y_4^{(2)\pm}(x)$ are shown for clearness. The evolution of this family with an increase in the parameter b to $b = -10$ is shown in Fig. 3, where a strong increase in the spread of all curves (by one order of magnitude) is seen. We note that the area under each curve is equal to unity (normalized rate).

4. We now determine the intervals $b^{(n)*} \le b < b^{(n-1)*}$ and construct families of curves $\gamma_k^{(n)}$ and $\lambda_k^{(n)}$ as functions of the parameter b for $b^{(n)*} \le b < 0$, where $n = 3, 4, \ldots$ is the interval number and k (number of modes) is a function of the interval number n. Figure 4 shows the fragment of the bifurcation pattern of flows for the third interval $b^{(3)*} \approx -176.5 \le b < b^{(2)*}$. The behavior of the family of curves $\gamma_k^{(3)}$ and $\lambda_k^{(3)}$ for $k = 5, 6, 7$ is similar to the behavior analyzed above for $n = 1$ and 2 and $k = 1, 2, 3$ and 3, 4, 5, respectively. Five-, six-, and seven-mode flows exist near the triple bifurcation point $b = b_*^{(3)} \approx -169.5$, and $\gamma_5^{(3)} = 0$ for $b = b_*^{(3)}$. The critical value $b = b^{(3)*}$ is the boundary of the interval and the turning (return) point of the curves $\gamma_7^{(2)}$ and $\lambda_7^{(2)}$. The curves $\gamma_k^{(3)}$ and $\lambda_k^{(3)}$ tend to the vertical asymptote for $b \longrightarrow -0$ more rapidly than the corresponding curves for $n = 1$ and 2. This pattern is qualitatively similar for any n value.

An analysis shows that the critical points $b_*^{(n)}$ for $\beta \longrightarrow 0$ ($0 < \beta \lesssim 10°$), where $\gamma_{2n-1}^{(n)} = 0$ and the bifurcations (triple points) of flows occur, are distributed according to

$$b_*^{(n)} = 6\pi n^2$$

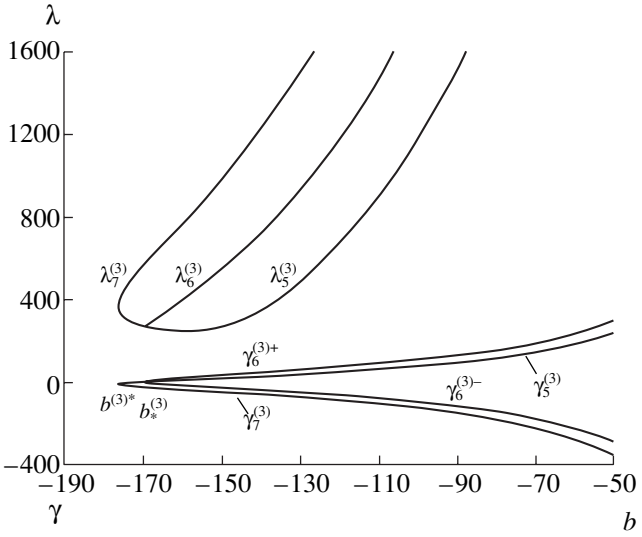


Fig. 4. Fragment of the bifurcation pattern for the flows $\gamma_{5,6,7}$ and $\lambda_{5,6,7}$ for the third Re interval $b^{(3)*} \le b < b^{(2)*}$.

(see the cases $n = 1, 2, 3$ above). The boundary points $b^{(n)*} < b_*^{(n)}$ are shifted to the left by a small value (about several units), which increases slightly with n. Thus, the distance between the points $b_*^{(n+1)}$ and $b_*^{(n)}$ is approximately equal to $6\pi(2n + 1)$, i.e., increases rapidly with n. The interval lengths $|b^{(n+1)*} - b^{(n)*}|$ are approximately equal to the above value.

It was found that k-mode steady flows for $k \ge 2n + 1$ may exist for $b < b_*^{(n)}$, where modes with $k \le 2n$ are absent. This property may be responsible for the loss of the stability of steady (laminar) flows and the beginning of the first stage of turbulization [3, 4, 6]. In particular, the most stable single-mode flows are impossible for $b < b_*^{(1)} \approx -18.8$ ($\approx 6\pi$). On the contrary, steady flows of all modes exist for $0 > b \ge -18.8$. We emphasize that odd $k = 2n + 1$ and even $k = 2n$ modes, $n = 1, 2, \ldots$, are symmetric and asymmetric, respectively, about the diffuser axis $x = \frac{1}{2}$ ($\theta = 0$). Even-mode flows exist in pairs. The only single-mode steady flow exists in the diffuser for $0 > b > -18.8 \approx 6\pi$, i.e., $0 > \mathrm{Re} > -54$ for the diffuser expansion angle $\beta = 10°$. This flow is symmetric.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 02-01-00252, 02-01-00157, and 03-01-96539) and the Council of the President of the Russian Federation for the Support of Young Russian Scientists and Leading Scientific Schools (project nos. NSh-1627.2003.1 and MK-2093.2003.01).

REFERENCES

1. G. B. Jeffery, Philos. Mag. Ser. 6 **29** (172), 455 (1915).
2. G. Hamel, Jahresber. Dtsch. Math. Ver. **25**, 34 (1917).
3. N. E. Kochin, I. A. Kibel', and N. V. Roze, *Theoretical Hydromechanics* (Fizmatgiz, Moscow, 1963), Vol. 2 [in Russian].
4. L. G. Loĭtsyanskiĭ, *Fluid Mechanics* (Nauka, Moscow, 1978) [in Russian].
5. L. D. Akulenko, D. V. Georgievskiĭ, and S. A. Kumakshev, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 1, 15 (2004).
6. G. K. Batchelor, *Introduction to Fluid Dynamics* (Cambridge Univ. Press, Cambridge, 1970; Mir, Moscow, 1973).
7. K. Millsaps and K. Pohlhausen, J. Aeronaut. Sci. **2** (3), 187 (1953).
8. L. D. Akulenko, D. V. Georgievskiĭ, and S. A. Kumakshev, Dokl. Akad. Nauk **383** (1), 46 (2002) [Dokl. Phys. **47**, 219 (2002)].
9. L. D. Akulenko and S. A. Kumashev, Dokl. Akad. Nauk **396** (4), 480 (2004) [Dokl. Phys. **49**, 361 (2004)].
10. L. D. Akulenko, S. A. Kumakshev, and S. V. Nesterov, Prikl. Mat. Mekh. **66**, 723 (2002).

Translated by R. Tyapaev

Doklady Physics, Vol. 49, No. 12, 2004, pp. 756–759. Translated from Doklady Akademii Nauk, Vol. 399, No. 5, 2004, pp. 625–629.

MECHANICS

Mathematical Simulation of the Intrusion of Sea Water into Coastal Freshwater Levels

É. N. Bereslavskiĭ

Presented by Academician G.G. Chernyĭ July 5, 2004

Received June 18, 2004

In previous works [1–7] on the hydrodynamic theory of the joint motion of fresh and salt water, it was assumed that fresh water flows in beds, the lower part of which is occupied by heavier salt water, whose unperturbed surface is always horizontal. The real initial contact surface of fluids is not necessarily horizontal, which is important for determining the position of the interface. This circumstance was first mentioned in [8], where the motion of ground water to a basin with a vertical bank was analyzed. This analysis was further developed in [9, 10], where the motion of water in a semi-infinite layer to a sea, as well as the case of the presence of a freshwater layer over the sea surface, was considered.

In this work, we generalize the above problems to the case of the percolation of fresh ground water in a finite rectangular bed to the sea with the freshwater layer over its surface. In the two types of flows under consideration, ground-water streams arrive at the sea from a side or from the bottom. By means of the derived exact analytical dependences and numerical calculations, the effect of each physical parameter of the models on the geometric and percolation characteristics is analyzed, and the features and character of percolation are studied. The solution derived in this work provides not only results for the above two limiting cases (for one particular set of the unknown parameters of the conformal mapping) but also the known results obtained by Polubarinova-Kochina for the classical problem of percolation in a rectangular dam [1, 2, 11]. For the case of lateral-inflow without the freshwater layer over the sea surface, the results for both schemes are compared for the same percolation parameters, and the dependence of the percolation characteristics on the initial position of the liquid interface is discussed.

1. We consider a plane flow of fresh ground water of density ρ_1 in a rectangular head bed of thickness T and width L, located on an impenetrable layer of halite, to the sea with salt water of density ρ_2 ($\rho_2 > \rho_1$). A freshwater layer is located over the saltwater level in the sea with the level t ($0 < t \le T$). Under the action of head H, the initially vertical interface between moving fresh water and immovable heavier salt water begins to deform in its right lower part, shifting to the left, towards the flow. After a certain time, the flow may become steady [12–14]; when the brine is stabilized, the interface becomes a streamline for fresh water, and the motion of sea water towards land takes the shape of a saltwater wedge (or tongue) penetrating into the freshwater bed (Fig. 1). This phenomenon is called the intrusion of salt water [5–7].

We assume that the flow of ground water satisfies Darcy's law and proceeds in the uniform isotropic soil. The soil, as well as the fluid percolated through it, is

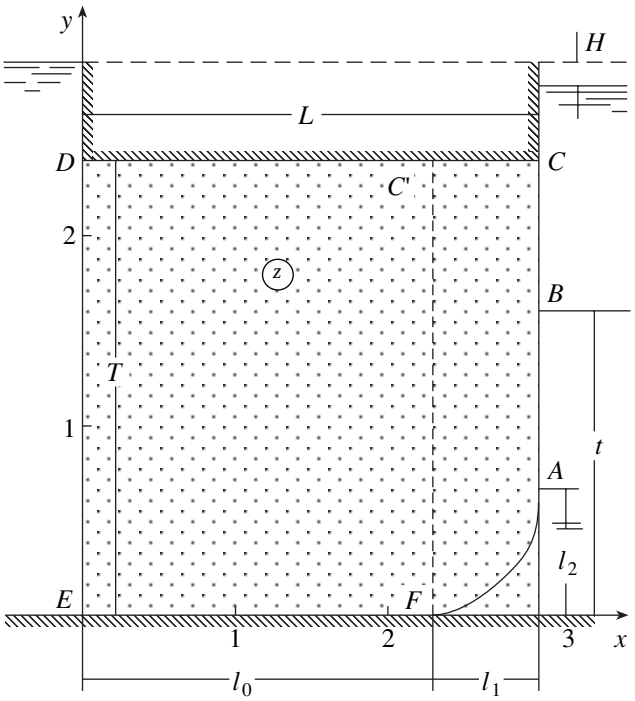


Fig. 1. Shape of the lateral-inflow flow calculated for T = 2.4, L = 3.0, ρ = 0.01, H = 0.036, and t = 1.5861.

State Academy of Civil Aviation,
St. Petersburg, 196210 Russia
e-mail: beres@nwgsm.ru

1028-3358/04/4912-0756$26.00

treated as incompressible. The effect of capillary and diffusion phenomena at the interface is ignored. Under these assumptions, which are traditional for the flows under consideration, the mathematical simulation of the percolation process under study reduces to determining the complex potential $\omega(z)$ under the boundary conditions

$$\begin{aligned} &AB\colon x = L, \quad \varphi = \rho(y-1); \\ &BC\colon x = L, \quad \varphi = 0, \quad \rho = \frac{\rho_2}{\rho_1} - 1; \\ &CD\colon y = T, \quad \psi = Q; \quad DE\colon x = 0, \quad \varphi = -H; \\ &EF\colon y = 0, \quad \psi = 0; \quad FA\colon \varphi = \rho(y-1), \quad \psi = 0. \end{aligned} \tag{1}$$

Here, φ and ψ are the velocity potential and stream function, respectively, that are divided by the soil hydraulic conductivity κ, and Q is the total percolation rate divided by κ. It is necessary to determine the position of the interface AF and, therefore, the width l_1 and height l_2 of the saltwater wedge intruded into the freshwater layer.

2. We introduce an auxiliary variable ζ and a function $z(\zeta)$ that conformally maps the upper ζ half-plane onto a z region (the correspondence of the points is shown in Fig. 2): $Z = \dfrac{dz}{d\zeta}$ and $F = \dfrac{d\omega}{d\zeta}$.

Using the Polubarinova-Kochina method [1, 2, 15], we obtain the parametric solution of the initial boundary value problem in the form

$$\begin{aligned} &Z = \frac{AiK'}{\Delta}, \quad F = \frac{Ai(K(\zeta) - iK'(\zeta))}{\Delta}, \\ &\Delta = \sqrt{(\zeta - c)(\zeta - d)(e - \zeta)}. \end{aligned} \tag{2}$$

Here, $K(\zeta)$ is the complete elliptic integral of the first kind as a function of $k^2 = \zeta$, $K'(\zeta) = K(1-\zeta)$, $k'^2 = 1 - \zeta$,

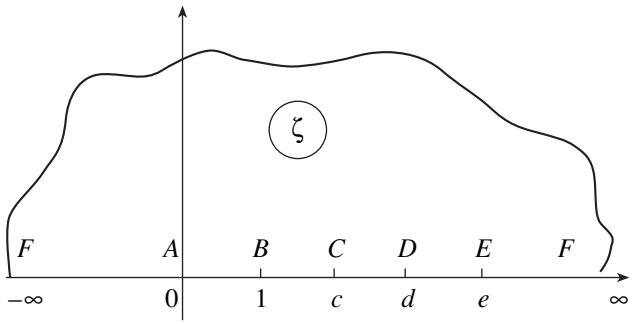


Fig. 2. Domain of the auxiliary parametric variable.

$A > 0$ is a scaling constant, and c, d, and e ($1 \le c \le d < e \le \infty$) are the parameters of the conformal mapping. Taking representation (2) for various sections of the boundary of the ζ region and integrating over the entire contour of the auxiliary region, we arrive at the parametric equations of the corresponding sections of the flow pattern. As a result, we obtain expressions for the geometric and percolation characteristics. Some of these expressions are used to find the unknown parameters of the conformal mapping, and others determine the dimensions of intrusion and rates.

3. Percolation through a rectangular dam. The Polubarinova-Kochina solution [1, 2, 11]. When points E and F are united in the z and ζ planes, i.e., for $e = \infty$, which corresponds to $l_0 = 0$, the flow region is mapped onto the mirror image of the motion region of ground water in the rectangular dam. The known solution of this problem, first obtained by Polubarinova-Kochina [Eqs. (10.19)–(10.28) on pp. 272, 273 in [1]], follows from Eq. (2) with $d = \dfrac{c}{c-1}$, $c = \dfrac{d}{d-1}$, and the replacement of c, d, L and l_1, l_2, $t - l_2$, T, $T - t$, Q_{AB}, and

Table 1. Values l_1, l_2, (above the lines) Q or (below the lines) L calculated for the given values T and (above the lines) L or (below the lines) Q. The values above and below the lines are obtained for the lateral- and bottom-inflow schemes, respectively

T	l_1	l_2	$\dfrac{Q}{L}$	$\dfrac{L}{Q}$	l_1	l_2	$\dfrac{Q}{L}$
1.0	0.3333 0.4462	0.3108 1.6826	0.0103 0.2268	2.000 0.024	0.6840 1.2152	0.6430 1.9595	0.0276 2.3861
1.6	0.6320 0.7278	0.5743 2.1905	0.0149 0.3548	3.000 0.036	1.1985 1.1755	1.0354 2.2029	0.0192 1.2399
2.0	0.8851 0.9281	0.7869 2.4246	0.0173 0.4332	4.000 0.048	1.7795 1.1529	1.3502 2.4537	0.0143 0.7017
2.4	1.1985 1.1461	1.0354 2.5812	0.0192 0.5029	5.000 0.060	2.3814 1.1436	1.5694 2.6491	0.0112 0.4087
3.0	1.9072 1.5438	1.5166 2.6465	0.0205 0.5783	6.000 0.072	2.8781 1.1402	1.6962 2.7747	0.0095 0.2518

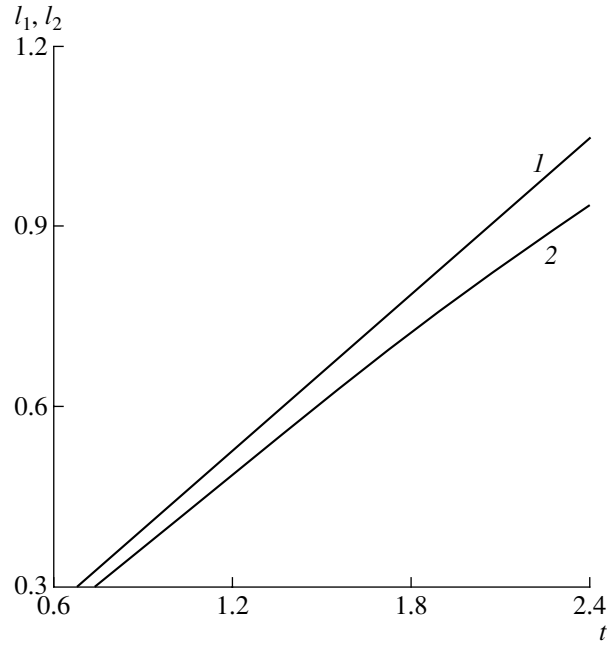


Fig. 3. Values l_1 and l_2 vs. t for $T = 2.4$, $L = 3.0$, $\rho = 0.01$, and $H = 0.036$.

Q_{BC} by a, b, l, $H_1 - H_2 - H_0$, H_0, H_1, H_2, Q_1, and Q_2, respectively.

4. Tables 1 and 2 present the results for the effect of determining the parameters T, L, ρ, and H on l_1, l_2, and Q for (above the lines) lateral inflow with $t = T$ and (below the lines) bottom inflow. In each of three blocks of the tables, one of the above parameters varies while the other parameters are fixed (basic variant): (lateral inflow) $T = 2.4$ and $L = 3.0$ and (bottom inflow) $\rho = 0.01$, $H = 0.036$, and $Q = 0.055$. Figure 3 shows (lines *1*) l_1 and (lines *2*) l_2 as functions of t, where $0 < t \le T$. An analysis of the tables and of Fig. 3 provides the following conclusions.

An increase in the dimensions of the bed, in the density of salt water, and in its thickness in the sea, as well as a decrease in the acting head, all increase the dimensions of intrusion. In this case, the relative dimensions of the wedge may be very large. For $T = 2.4$, $L = 6.0$, and $\rho = 0.01$, we have $l_2 = 1.6962$, and $l_1 = 2.1462$ for $T = 2.4$, $L = 3.0$, and $\rho = 0.012$; i.e., the height and width of the wedge may reach 70.6 and 71.5% of the thickness and width of the bed, respectively. It is remarkable that $l_1 \approx l_2$ for small T, L, and ρ and large H, whereas $l_1 \approx ml_2$, where $1.1 \le m \le 1.7$ for large T, L, and ρ and small H.

The dimensions of intrusion as functions of the thickness of the saltwater layer in the sea are of special interest. As is seen in Fig. 3, l_1 and l_2 depend linearly on t. For the basic variant, $l_1 = 0.44t - 0.1$ and $l_2 = 0.4t$ can be taken. In this case, the width of the wedge is larger than its height by 8–12%. At the same time, when the saltwater level in the sea increases, the ratio $\frac{l_1}{t}$ remains virtually unchanged (43.6%) with a tendency to a slight decrease, whereas the ratio $\frac{l_2}{t}$ decreases from 40.5% for $t = 0.6893$ to 38.9% for $t = T$. The latter property means that the relative dimensions of the saltwater wedge and, therefore, the degree of intrusion increases with the thickness of fresh water over salt water.

5. Figure 4 shows the traditional scheme of the bottom inflow of fresh ground water into the sea from the coastal head aquifer. In this case, the problem is reduced to the determination of the function $\omega(z)$ under boundary conditions (1), where conditions on the sections AB and BC are replaced by the conditions $y = T$

Table 2. Values l_1, l_2, (above the lines) Q or (below the lines) L calculated for the given values ρ and H. The values above and below the lines are obtained for the lateral- and bottom-inflow schemes, respectively

ρ	l_1	l_2	$\frac{Q}{L}$	H	l_1	l_2	$\frac{Q}{L}$
0.004	0.1237 1.0700	0.1234 4.1026	0.0215 0.5450	0.032	1.6816 1.1768	1.2991 2.4465	0.0151 0.3568
0.006	0.3925 1.0849	0.3852 3.4582	0.0212 0.5362	0.048	0.6571 1.1025	0.6255 2.8319	0.0276 0.9902
0.008	0.7550 1.1087	0.7084 2.9802	0.0205 0.5227	0.064	0.3689 1.0928	0.3524 3.0946	0.0365 1.6846
0.010	1.1985 1.1461	1.0354 2.5812	0.0192 0.5029	0.080	0.1930 1.0829	0.1919 3.3484	0.0467 2.3854
0.012	2.1462 1.2104	1.4697 2.2121	0.0152 0.4729	0.096	0.0721 1.0765	0.0720 3.5807	0.0614 3.0879

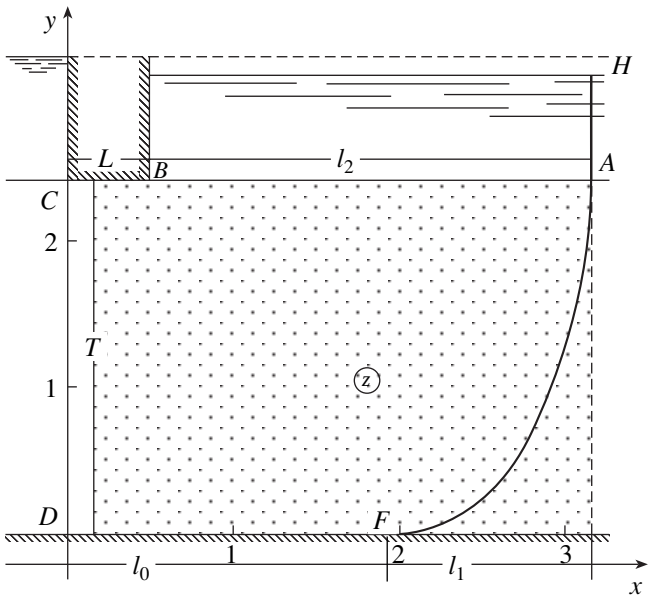


Fig. 4. Shape of the bottom-inflow flow calculated for $T = 2.4$, $\rho = 0.01$, $H = 0.036$, and $Q = 0.055$.

and $\varphi = 0$, and l_2 is the length of the section of the inset of fresh water into the sea bottom. The parametric solution has the form

$$Z = -\frac{2A\arcsin\sqrt{1-\zeta}}{\pi\Delta}, \quad F = \frac{A\rho i}{\Delta}, \qquad \Delta = \sqrt{\zeta(1-\zeta)(d-\zeta)(e-\zeta)}. \tag{3}$$

In this case, to determine the unknown parameters, it is necessary to solve the equation $\frac{K'}{K} = \frac{Q}{H}$, along with other equations. This equation restricts the physical parameters Q and H and, therefore, the applicability domain of the chosen flow scheme. For this reason, in contrast to the lateral-inflow scheme, when performing numerical calculations it is more convenient to specify the rate Q, which is a key parameter in solving the last equation, rather than the bed width L that is determined later.

6. According to Tables 1 and 2 (data below the lines), an increase in the thickness of the bed leads to the increase in the dimensions of the wedge; i.e., this dependence is similar to that in the above problem. The dependence of the width l_1 on T is qualitatively similar to the lateral-inflow case, and l_1 depends linearly on ρ and H. However, the dependence of l_2 on ρ and H changes strongly compared to the first scheme: its increase is caused by a decrease in the saltwater density and an increase in the acting head and rate, which is in turn associated with a decrease in the bed width.

Calculations show that the inequalities

$$0.4 \le \frac{l_1^{(2)}}{l_1^{(1)}} \le 15, \quad 1.5 \le \frac{l_2^{(2)}}{l_2^{(1)}} \le 50,$$

where the superscript indicates the scheme of calculations, are satisfied for the parameters presented in Tables 1 and 2. The latter inequality means that l_2 in the problem for the bottom inflow is larger than the height of the wedge in the problem for the lateral inflow by a factor of at least 1.5.

When points D and E are united in the z and ζ planes, which corresponds to $L = \infty$ and $d = e$, we arrive at the flow in the semi-infinite head aquifer [9]. The results for this case are obtained from Eq. (4) with $k = 1$ and $k' = 0$.

REFERENCES

1. P. Ya. Polubarinova-Kochina, *Theory of the Motion of Ground Water* (Gostekhizdat, Moscow, 1952) [in Russian].
2. P. Ya. Kochina, *Hydrodynamics and Percolation Theory: Selected Works* (Nauka, Moscow, 1991) [in Russian].
3. *Development of Investigations on Percolation Theory in the USSR (1917–1967)* (Nauka, Moscow, 1969) [in Russian].
4. G. K. Mikhaĭlov and V. N. Nikolaevskiĭ, in *Mechanics in the USSR for 50 Years* (Nauka, Moscow, 1970), Vol. 2, pp. 585–648 [in Russian].
5. J. Bear, D. Saslavsky, and S. Irmay, *Physical Principles of Water Percolation and Seepage* (UNESCO, Paris, 1968; Mir, Moscow, 1971).
6. A. P. Frolov and M. G. Khublaryan, Vodn. Resur., No. 2, 58 (1986).
7. M. G. Khublaryan and A. P. Frolov, *Simulation of Intrusion Processes in Reaches and Ground Aquifers* (Nauka, Moscow, 1989) [in Russian].
8. G. K. Mikhaĭlov, Dokl. Akad. Nauk SSSR **110** (6), 945 (1956).
9. É. N. Bereslavskiĭ, Izv. Akad. Nauk, Ser. Mekh. Zhidk. Gaza, No. 3, 96 (2003).
10. É. N. Bereslavskiĭ, Prikl. Mat. Mekh. **67**, 836 (2003).
11. P. Ya. Polubarinova-Kochina, Prikl. Mat. Mekh. **4**, 53 (1940).
12. V. I. Aravin and S. N. Numerov, *Theory of the Motion of Fluids and Gases in Undeformed Porous Media* (Gostekhizdat, Moscow, 1953) [in Russian].
13. P. Ya. Polubarinova-Kochina, *Some Problems of the Plane Motion of Ground Water* (Akad. Nauk SSSR, Moscow, 1942) [in Russian].
14. P. Ya. Polubarinova-Kochina, Prikl. Mat. Mekh. **4**, 101 (1940).
15. P. Ya. Kochina, É. N. Bereslavskiĭ, and N. N. Kochina, Preprint No. 567, IPM RAN (Inst. for Problems in Mechanics, Russian Academy of Sciences, 1996).

Translated by R. Tyapaev

Doklady Physics, Vol. 49, No. 12, 2004, pp. 760–762. Translated from Doklady Akademii Nauk, Vol. 399, No. 5, 2004, pp. 630–632.

MECHANICS

The Duality Theorem in Nonlocal Fluid Dynamics

A. A. Rodionov

Presented by Academician S.N. Kovalev March 26, 2004

Received April 16, 2004

In classical fluid dynamics, the duality theorem was proven by H.A. Lorentz for slow steady-state flows of incompressible Newtonian fluids [1]. In [2], the theorem was generalized to the case of a flow of two Newtonian fluids having different viscosities (μ_1 and μ_2) and was applied for describing flows of "pure" and dispersed media. The theorem is formulated in the following manner.

Theorem 1. *For two different steady-state Stokes flows of incompressible Newtonian fluid, the following relation holds*:

$$\mu_2\int\limits_S \partial\mathbf{S}\cdot\mathbf{P}^{(1)}\cdot\mathbf{u}^{(2)} = \mu_1\int\limits_S \partial\mathbf{S}\cdot\mathbf{P}^{(2)}\cdot\mathbf{u}^{(1)}, \qquad (1)$$

where S is an arbitrary closed surface drawn in the fluid, $\mathbf{u}^{(1)}$ *and* $\mathbf{u}^{(2)}$ *are the velocities of medium motion in the flows, and* $\mathbf{P}^{(1)}$ *and* $\mathbf{P}^{(2)}$ *are the stress tensors.*

For qualitative and quantitative analysis of the regularities of medium flows, we derive an analog of the generalized duality theorem in the framework of nonlocal theory.

For incompressible continuum media, we consider two flows specified as ($\mathbf{u}^{(1)}$, $\mathbf{P}^{(1)}$) and ($\mathbf{u}^{(2)}$, $\mathbf{P}^{(2)}$), which are described by the equations

$$\nabla\cdot\mathbf{P}^{(1)} = \sum_{\alpha=1}^{3}\frac{\partial P^{(1)}_{\alpha\beta}}{\partial r_\alpha} = 0, \quad \beta = 1, 2, 3, \quad l = 1, 2. \quad (2)$$

In nonlocal fluid dynamics, the quantities $\mathbf{P}^{(l)}$ have the form [3, 4]

$$P^{(1)}_{\alpha\beta} = -\delta_{\alpha\beta}p^{(l)} + 2\int\limits_V \partial r'_1\partial r'_2\partial r'_3 L_{(l)}(\mathbf{r}, \mathbf{r}'; \alpha, \beta)E^{(1)}_{\alpha\beta}(\mathbf{r}'), \qquad (3)$$

where $\mathbf{r}$ is the spatial Cartesian coordinates; $p^{(l)}$ is pressure; $\mathbf{E}^{(l)}$ is the strain-rate tensor,

$$E^{(l)}_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial u^{(l)}_\beta}{\partial r_\alpha} + \frac{\partial u^{(l)}_\alpha}{\partial r_\beta}\right) = E^{(l)}_{\beta\alpha}; \qquad (4)$$

V is the volume occupied by the medium flow; and $L^{(l)}$ are the kernels describing momentum transfer and taking into account nonlocal interrelations.

Further, we multiply $P^{(1)}_{\alpha\beta}$ by $E^{(2)}_{\alpha\beta}$, perform summation, and arrive at

$$P^{(1)}_{\alpha\beta}E^{(2)}_{\alpha\beta} = \mathbf{P}^{(1)}:\mathbf{E}^{(2)} = 2\int\limits_V \partial r'_1\partial r'_2\partial r'_3 L_{(l)}E^{(1)}_{\alpha\beta}(\mathbf{r}')E^{(2)}_{\alpha\beta}(\mathbf{r}). \qquad (5)$$

Hereafter, summation over dummy indices from one to three is implied. Performing identity transformations of ($\mathbf{P}^{(1)}:\mathbf{E}^{(2)}$), we obtain

$$\mathbf{P}^{(1)}:\mathbf{E}^{(2)} = P^{(1)}_{\alpha\beta}\frac{\partial u^{(2)}_\beta}{\partial r_\alpha} - P^{(1)}_{\alpha\beta}\Omega^{(2)}_{\alpha\beta}, \qquad (6)$$

where $\Omega^{(l)}_{\alpha\beta} = \dfrac{\partial u^{(l)}_\beta}{\partial r_\alpha} - E^{(l)}_{\alpha\beta}$.

Further, the formula

$$\frac{\partial}{\partial r_\alpha}(P^{(1)}_{\alpha\beta}u^{(2)}_\beta) = u^{(2)}_\beta\frac{\partial P^{(1)}_{\alpha\beta}}{\partial r_\alpha} + P^{(1)}_{\alpha\beta}\frac{\partial u^{(2)}_\beta}{\partial r_\alpha} = P^{(1)}_{\alpha\beta}\frac{\partial u^{(2)}_\beta}{\partial r_\alpha}, \qquad (7)$$

where the validity of relation (2) is taken into account, is applied to relation (6). Upon bringing together formulas (5)–(7), we multiply the result by $\partial r_1\partial r_2\partial r_3$ and integrate it over an arbitrary volume V_r contained in the region of the flow and bounded by the surface S_r. As a result, we have

$$2\int\limits_{V_r}\partial r_1\partial r_2\partial r_3\int\limits_V \partial r'_1\partial r'_2\partial r'_3 L_{(l)}E^{(1)}_{\alpha\beta}(\mathbf{r}')E^{(2)}_{\alpha\beta}(\mathbf{r}) + \int\limits_{V_r}\partial\mathbf{r}P^{(1)}_{\alpha\beta}\Omega^{(2)}_{\alpha\beta} = \int\limits_{V_r}\partial\mathbf{r}\frac{\partial}{\partial r_\alpha}(P^{(1)}_{\alpha\beta}u^{(2)}_\beta). \qquad (8)$$

St. Petersburg Research Center,
Russian Academy of Sciences,
Universitetskaya nab. 15, St. Petersburg, 199034 Russia

1028-3358/04/4912-0760$26.00

Further, the divergence theorem, in the form

$$\int_{V_r} \partial r_1' \partial r_2' \partial r_3' \frac{\partial}{\partial r_\alpha} (P^{(1)}_{\alpha\beta} u^{(2)}_\beta) = \int_{S_r} \partial \mathbf{S} \cdot \mathbf{P}^{(1)} \cdot \mathbf{u}^{(2)}, \quad (9)$$

is applied to the expression on the right-hand side of (8).

Upon mutual interchanging of the superscripts 1 and 2 and dividing the result by (8), we ultimately obtain

$$M^{(2)}_\Omega \int_{S_r} \partial \mathbf{S} \cdot \mathbf{P}^{(1)} \cdot \mathbf{u}^{(2)} = M^{(1)}_\Omega \int_{S_r} \partial \mathbf{S} \cdot \mathbf{P}^{(2)} \cdot \mathbf{u}^{(1)}, \quad (10)$$

where

$$M^{(1)}_\Omega = 2 \int_{V_r} \partial \mathbf{r} \int_V \partial \mathbf{r}' L_{(1)} E^{(1)}_{\alpha\beta}(\mathbf{r}') E^{(2)}_{\alpha\beta}(\mathbf{r}) + \int_{V_r} \partial \mathbf{r} P^{(1)}_{\alpha\beta} \Omega^{(2)}_{\alpha\beta},$$

$$M^{(2)}_\Omega = 2 \int_{V_r} \partial \mathbf{r} \int_V \partial \mathbf{r}' L_{(2)} E^{(2)}_{\alpha\beta}(\mathbf{r}') E^{(1)}_{\alpha\beta}(\mathbf{r}) + \int_{V_r} \partial \mathbf{r} P^{(2)}_{\alpha\beta} \Omega^{(1)}_{\alpha\beta}.$$

The resulting formula is valid in the general case of an asymmetric stress tensor, $P^{(l)}_{\alpha\beta} \neq P^{(l)}_{\beta\alpha}$ ($l = 1, 2$). If $P^{(l)}_{\alpha\beta} = P^{(l)}_{\beta\alpha}$ ($l = 1, 2$), relation (10) reduces to the form

$$M^{(2)} \int_{S_r} \partial \mathbf{S} \cdot \mathbf{P}^{(1)} \cdot \mathbf{u}^{(2)} = M^{(1)} \int_{S_r} \partial \mathbf{S} \cdot \mathbf{P}^{(2)} \cdot \mathbf{u}^{(1)}, \quad (11)$$

where

$$M^{(1)} = \int_{V_r} \partial \mathbf{r} \int_V \partial \mathbf{r}' L_{(1)} E^{(1)}_{\alpha\beta}(\mathbf{r}') E^{(2)}_{\alpha\beta}(\mathbf{r}),$$

$$M^{(2)} = \int_{V_r} \partial \mathbf{r} \int_V \partial \mathbf{r}' L_{(2)} E^{(2)}_{\alpha\beta}(\mathbf{r}') E^{(1)}_{\alpha\beta}(\mathbf{r}).$$

Within nonlocal fluid dynamics, relations (10) and (11) represent the duality theorem for the cases of an asymmetric and a symmetric stress tensor, respectively. These relations hold for arbitrary transfer kernels, in particular, for kernels belonging to the $\{L\}_\delta$ class [3, 4], which reduce to δ-functions in the "local" limit of a Newtonian fluid.

We now suppose that $L_{(l)} \in \{L\}_\delta$, $l = 1, 2$. As the parameters characterizing the nonlocality of the integral kernels tend to zero, $\varepsilon^{(1)}, \varepsilon^{(2)} \longrightarrow 0$, formula (11) transforms into the relation corresponding to the generalized duality theorem [2], namely,

$$\mu^{(2)}_0 \int_{S_r} \partial \mathbf{S} \cdot \mathbf{P}^{(1)}_0 \cdot \mathbf{u}^{(2)}_0 = \mu^{(1)}_0 \int_{S_r} \partial \mathbf{S} \cdot \mathbf{P}^{(1)}_0 \cdot \mathbf{u}^{(1)}_0, \quad (12)$$

where $\mu^{(l)}_0$, $l = 1, 2$, are the coefficients of viscosity of continuum media and zeros in the subscript labels are the respective limiting values for Newtonian fluid.

Let us consider the corollaries from the duality theorem.

First, we apply it to analysis of the motion of an incompressible medium making allowance for nonlocality effects. We consider a plane Poiseuille flow under the effect of a constant pressure gradient between two infinite parallel plates separated by the distance $2R$. For one of the flows, we take a flow of Newtonian fluid characterized by a viscosity coefficient μ_0, while, for the second flow, we consider a flow of a medium featuring manifestations of nonlocality effects. Let S_r be the surface of a rectangular parallelepiped formed by $1 \times 2R \times 1$ edges and oriented in such a way that its bases touch the plates and that the flow is orthogonal to one of its $1 \times 2R$ faces. We choose the r_2 axis to be aligned along the flow at the distance R from the plates and the r_3 axis to be orthogonal to the plates. In this case, relation (11) takes the form

$$\frac{\partial p^{(2)}}{\partial r_2} Q^{(1)} \Big/ \frac{\partial p^{(1)}}{\partial r_2} Q^{(2)} = \frac{\int_{-R}^{R} \partial r_3 \int_{-R}^{R} \partial r_3' L(r_3, r_3') \frac{\partial u_2^{(2)}}{\partial r_3'} \frac{\partial u_2^{(1)}}{\partial r_3}}{\mu_0 \int_{-R}^{R} \partial r_3 \frac{\partial u_2^{(2)}}{\partial r_3} \frac{\partial u_2^{(1)}}{\partial r_3}}, \quad (13)$$

where $Q^{(l)}$ is the medium flow rate ($l = 1, 2$).

We denote the right-hand side of formula (13) by μ^P; thus,

$$\mu^P = \frac{\int_{-R}^{R} \partial r_3 \int_{-R}^{R} \partial r_3' L(r_3, r_3') \frac{\partial u_2}{\partial r_3'} r_3}{\int_{-R}^{R} \partial r_3 \frac{\partial u_2}{\partial r_3} r_3}. \quad (14)$$

Here, we allow for the fact that the quantity $\frac{\partial u_2^{(1)}}{\partial r_3}$ is proportional to r_3 and the superscript on the velocity of the structured medium is omitted. Taking into account the relationship between $\frac{\partial p^{(1)}}{\partial r_2}$ and $Q^{(1)}$, we can then write

$$Q^{(2)} = -\frac{2}{3} \frac{R^3}{\mu^P} \frac{\partial p^{(2)}}{\partial r_2}. \quad (15)$$

On the basis of this formula, we conclude that μ^P is the coefficient of the effective viscosity of a medium for Poiseuille flow, this coefficient being defined as the proportionality factor in relation (15).

We now consider a Couette flow between two infinite parallel plates moving in opposite directions at velocity u. By applying the duality theorem to this Couette flow in the same way as was done above for the Poiseuille flow, we obtain

$$\frac{\mathbf{P}^{(2)}(R)u}{\mathbf{P}^{(1)}(R)u} = \frac{\int_{-R}^{R}\partial r_3\int_{-R}^{R}\partial r_3' L(r_3, r_3')\dfrac{\partial u_2^{(2)}}{\partial r_3'}\dfrac{\partial u_2^{(1)}}{\partial r_3}}{\mu_0\int_{-R}^{R}\partial r_3\dfrac{\partial u_2^{(2)}}{\partial r_3}\dfrac{\partial u_2^{(1)}}{\partial r_3}}, \quad (16)$$

where $\mathbf{P}^{(l)}(R)$ is the value of the stress tensor at the boundary ($l = 1, 2$).

Further, we denote the right-hand side of (16) by μ^C:

$$\mu^C = \frac{\int_{-R}^{R}\partial r_3\int_{-R}^{R}\partial r_3' L(r_3, r_3')\dfrac{\partial u_2}{\partial r_3'}}{\int_{-R}^{R}\partial r_3\dfrac{\partial u_2}{\partial r_3}}. \quad (17)$$

In Eq. (17), we have omitted the superscript standing for the velocity of the structured medium and have considered that $\dfrac{\partial u_2^{(1)}}{\partial r_3} = \text{const}$.

We now rewrite Eq. (16) in the form

$$\mathbf{P}^{(2)}(R) = \frac{\mu^C u}{R}. \quad (18)$$

It can be concluded that μ^C is the coefficient of the effective viscosity of a medium in the case of a Couette flow.

Combining (14) and (17) into one expression, we have

$$\mu^{(m)} = \frac{\int_{-R}^{R}\partial r_3\int_{-R}^{R}\partial r_3' L(r_3, r_3')\dfrac{\partial u_2}{\partial r_3'}r_3^m}{\int_{-R}^{R}\partial r_3\dfrac{\partial u_2}{\partial r_3}r_3^m}, \quad (19)$$

where $m = 0$ for Couette flow, $m = 1$ for Poiseuille flow, and the superscript m not enclosed by parentheses stands for a conventional power-law exponent.

By definition, $\mu^{(m)}$ is a functional of the external-flow velocity gradient, of the geometry of the flow boundaries, and of the momentum-transfer kernel, this functional taking into account nonlocality effects. Experimental investigations performed under various external conditions for flows of media having an internal structure confirm this conclusion. In particular, it is known that, in experimentally determining rheological properties of heterogeneous mixtures in viscometers of the Couette and Poiseuille types, one obtains different values of $\mu^{(m)}$ for preassigned flow modes [2].

In the general case, definition (19) makes it possible, on the basis of the constructed models of the transfer kernels for various media featuring nonlocality effects, to trace the effect of the above factors on the value of the effective-viscosity coefficient for specific flow modes.

Thus, the duality theorem has been proven here in the framework of nonlocal fluid dynamics. It has been shown that, in the local limit, the basic relations reduce to the duality theorem for the motion of two Newtonian fluids having different viscosities.

For a medium having an internal structure, the duality theorem makes it possible to introduce the effective-viscosity coefficient as an explicit functional of the velocity gradient, the geometry of the flow boundaries, and the momentum-transfer kernel. A generalization of the Hagen–Poiseuille law (Poiseuille flow) and of equations that relate the characteristics of Couette flow to the case of fluid motions involving nonlocality effects has been obtained as a corollary of the theorem.

REFERENCES

1. H. A. Lorentz, Abh. Theor. Phys. **1**–**3** (1906).
2. J. Happel and H. Brenner, *Low Reynolds Number Hydrodynamics* (Prentice-Hall, Englewood Cliffs, 1965; Mir, Moscow, 1976).
3. B. V. Filippov and T. A. Khantuleva, *Boundary Value Problems in Nonlocal Fluid Dynamics* (Izd. Leningr. Gos. Univ., Leningrad, 1984).
4. A. A. Rodionov and T. A. Khantuleva, in *Dynamics of Processes in Gases and Solids*, Ser.: *Phys. Mechanics* (Leningrad, 1984), No. 5.

Translated by T. Syromyatnikova

Doklady Physics, Vol. 49, No. 12, 2004, pp. 763–767. Translated from Doklady Akademii Nauk, Vol. 399, No. 5, 2004, pp. 633–637.

MECHANICS

Stabilization of a Hypersonic Boundary Layer by Ultrasound-Absorbing Coatings with a Regular Microstructure

Corresponding Member of the RAS V. M. Fomin[1], A. V. Fedorov[2], V. F. Kozlov[2], A. N. Shiplyuk[1], A. A. Maslov[1,*], E. V. Burov[1], and N. D. Malmuth[3]

Received April 26, 2004

INTRODUCTION

In the presence of small perturbations of the free flow (which are characteristic of flight conditions), the initial stage of the laminar–turbulent transition is characterized by exciting unstable modes and their enhancement to the nonlinear phase, terminating in a transition to the turbulent flow [1]. Stability theory and experiments show that perturbations of the first and second modes dominate in a hypersonic boundary layer. The cooling of a wall, which naturally occurs on the surface of a hypersonic aircraft with a heat-protective coating, stabilizes the first mode [2] and destabilizes the second mode [3].

The second mode belongs to a family of channel acoustic modes propagating in a waveguide between the wall and acoustic line [1]. The second mode is the dominant instability of the boundary layer for hypersonic aircrafts of a predominantly planar shape with a sharp leading edge and large local Mach numbers ($M_e > 6$).

The second mode in fast flows is associated with perturbations of a relatively high ultrasonic frequency. It was assumed in [4] that a passive ultrasound-absorbing coating (UAC) with a porous microstructure can damp these perturbations and does not lead to a premature transition due to the roughness effect. Theoretical analysis in the framework of the linear theory of stability in inviscid [4] and viscous [5] approximations showed that a relatively thin coating (with a thickness of about one-tenth the thickness of the boundary-layer displacement) can strongly reduce the increment of the second-mode growth. Those works were devoted to analyzing the UAC in the form of a porous layer with equidistant blind cylindrical microholes.

Experiments [6] carried out in a shock tube on a sharp cone for Mach numbers M = 5–6 qualitatively corroborated the theoretical conclusions. It was found that the flow remained laminar over the entire length of the model on the surface perforated by blind cylindrical holes (approximately 100 holes per 1 mm^2), whereas the transition on the smooth surface occurs in the middle part of the cone. Since the characteristics of perturbations in the boundary layer were not measured, the experiments did not provide any direct evidence of the presence of the second mode, its stabilization by the porous coating, or its role in the transition process.

The stability of the boundary layer on the sharp cone that had a half-angle of 7° and was coated with a UAC with a chaotic microstructure (metallic batting) [7] was analyzed in a T-326 wind tunnel at the Institute of Theoretical and Applied Mechanics, Siberian Division, Russian Academy of Sciences, for a Mach number M = 5.95. This analysis showed that the porous coating stabilized the second mode and slightly destabilized the first mode. A theoretical model of the absorption of acoustic energy on a chaotic-structure porous wall was developed in the framework of the classical approach. The absorption coefficient was calculated using laboratory measurements on the absorption of sound by samples of the coating. Good agreement between the theoretical increments of the growth of perturbations and the experimental data was achieved for two-dimensional perturbations of the second mode. However, the theoretical model was semiempirical, because direct calculation of the acoustic properties of the coating (without the use of experimental data) was impossible due to its complex microstructure.

This work is focused on theoretical and experimental investigations of the stability of the boundary layer on the surface with the regular-structure UAC. The the-

[1] *Institute of Theoretical and Applied Mechanics, Siberian Division, Russian Academy of Sciences, ul. Institutskaya 4/1, Novosibirsk, 630090 Russia*

** e-mail: maslov@itam.nsc.ru*

[2] *Faculty of Aeromechanics and Aircraft Engineering, Moscow Institute of Physics and Technology, ul. Gagarina 16, Zhukovskiĭ, Moscow oblast, 140180 Russia*

[3] *Rockwell Scientific Company, 1049 Camino Dos Rios, Thousand Oaks, California, USA*

1028-3358/04/4912-0763$26.00 © 2004 MAIK “Nauka/Interperiodica”

oretical investigations are aimed at both developing a closed model of the effect of the UAC on the development of perturbations in the rarefied gas and acquiring quantitative data on the characteristics of perturbations. The experimental investigations were aimed at determining the detailed wave characteristics of perturbations in order to verify the theoretical model.

POROUS COATING

The porous coating was made of a 0.45-mm-thick stainless-steel sheet perforated by the laser-drilling method. In the sheet closely attached to the smooth cone surface, the pores were blind cylindrical holes that had a diameter of 50 μm, were spaced by 100 μm, and had a total open area (porosity) of 20%. The characteristics of the porous coating are similar to those used in experiments reported in [6].

EXPERIMENTS

The experiments were carried out in the T-326 wind tunnel at the Institute of Theoretical and Applied Mechanics, Siberian Division, Russian Academy of Sciences. The free flow was characterized by a deceleration temperature of T_0 = 385–400 K and a Mach number of M_∞ = 5.95; the single Reynolds number was $Re_{1\infty} = (11.5–12.3) \times 10^6$ m^{-1} and the temperature surface factor was $\frac{T_w}{T_0}$ = 0.80–0.84, where T_w is the model wall temperature.

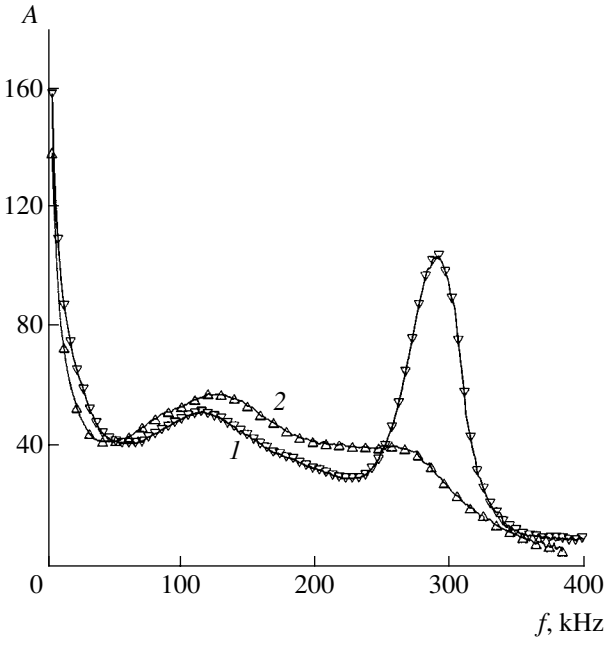


Fig. 1. Amplitude–frequency spectra of mass-rate pulsations on the (*1*) smooth and (*2*) porous surfaces for $Re_{ex} = 4.5 \times 10^6$.

The model under investigation was a 0.5-m-long sharp cone with a half-angle of 7°. It was placed in a flow at zero angle of attack. One half of the cone surface (between generatrices) was coated with the UAC, and the other half remained smooth. The leading boundary of the UAC was located at a distance of 182 mm from the vertex of the model, and the length of the coating area was equal to 263 mm. Artificial perturbations were introduced into the flow by means of an rf electric glow discharge initiated in a chamber placed within the model. The perturbations penetrated into the boundary layer through a hole that had a diameter of 0.4 mm and was spaced by 69 mm from the vertex of the model. The average and pulsation characteristics of the boundary layer were measured in the longitudinal (X), azimuthal (Θ), and normal (Y) directions to the cone surface by a heat loss anemometer. The results of the measurements were subjected to Fourier analysis, which provided the wave characteristics of the pulsations. Such a procedure had been successfully applied to analyze the development of wave packets in hypersonic boundary layers [8].

Measurements showed that the flow remained laminar on both the smooth and porous surfaces. The mean characteristics of the flow in the boundary layer (velocity profile, boundary layer thickness, etc.) and the distributions of integral pulsations of the mass rate were close to each other on both sides of the model. Therefore, the porous surface does not change the mean characteristics of the flow or the rms pulsations of the mass rate. The spectra of the mass-rate pulsations were measured for the Y coordinate corresponding to the maximum of pulsations. Figure 1 shows the amplitude–frequency spectra of perturbations on the (*1*) smooth and (*2*) porous surfaces in the X section corresponding to the maximum of the second-mode amplitude on the surface with the UAC ($Re_{ex} = 4.5 \times 10^6$). It is seen that the UAC efficiently suppresses perturbations of the second mode (the maximum of the amplitude on the porous wall is approximately equal to one third of the value on the smooth wall) and virtually does not affect the first mode.

Artificial perturbations with a frequency of 275 kHz were generated in the boundary layer. The second mode has its maximum amplitude at this frequency in the spectra of natural perturbations measured on the smooth wall. Figure 2 shows the amplitude wave spectra for artificial-perturbation packets on the porous side as functions of the transverse wavenumber β. The entire wave packet is located in the range $\beta = \pm 0.5$, which approximately corresponds to the range ±20° of the wave-vector slope. The amplitude is maximal for plane waves ($\beta = 0$). The dimensionless longitudinal phase velocity $C_x = 0.9$ is virtually independent of any increase in the coordinate X and is the same for the smooth and porous surfaces. These data show that the wave packet consists predominantly of second-mode perturbations. Growth curves for natural and artificial

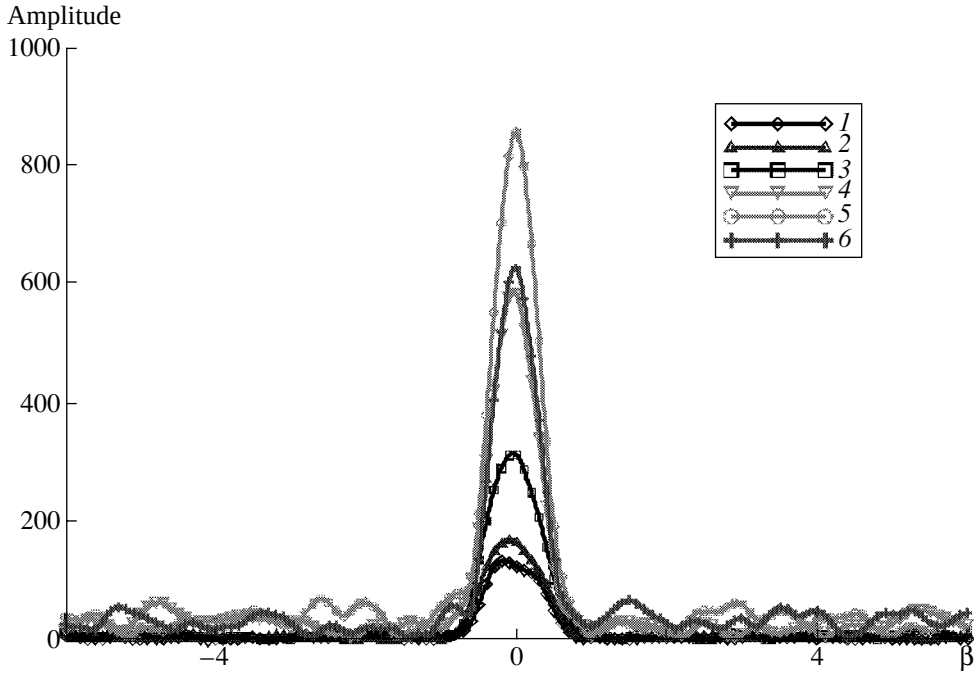


Fig. 2. β spectra of the amplitude of artificial perturbations in an artificial wave packet on the side with the ultrasound-absorbing coating for Re_{ex} = (*1*) 2.22×10^6, (*2*) 2.77×10^6, (*3*) 3.35×10^6, (*4*) 3.91×10^6, (*5*) 4.47×10^6, and (*6*) 5.05×10^6.

perturbations of the mass rate on the smooth and porous walls are shown in Fig. 4 (the amplitude of artificial perturbations corresponds to $\beta = 0$). The data are given for a frequency of 275 kHz. The growth curves virtually coincide on the porous wall. Therefore, natural perturbations at this frequency are predominantly two-dimensional waves of the second mode.

THEORY AND COMPARISON WITH EXPERIMENT

The problem of linear stability under three-dimensional perturbations in the two-dimensional boundary layer on the porous wall is formulated in terms of the asymptotic method of many scales [9]. The analysis takes into account that the middle flow is nonparallel because the boundary-layer thickness increases downstream. Effects associated with the curvature of the cone surface and the cone divergence of streamlines are ignored. In contrast to the classical problem of the stability of the boundary layer on the smooth wall, boundary conditions for perturbations on the porous surface are imposed with allowance for the acoustic conductivity of the porous coating, which depends on the microstructure of the UAC. As was shown in [5, 10], the amplitude functions of wave perturbations of the form

$$(u, v, w, p, \theta)(y)\exp[i(\alpha x + \beta z - i\omega t)]$$

satisfy the boundary conditions

$$u(0) = 0, \quad w(0) = 0, \quad v(0) = A_y p(0), \quad \theta(0) = 0. \quad (1)$$

Here, u, w, and v are the amplitudes of perturbations in the longitudinal, transverse, and normal velocity components, respectively, divided by the middle-flow velocity U_e^*; θ is the temperature-perturbation amplitude divided by T_e^*; and p is the pressure amplitude divided by the double velocity head $\rho_e^* U_e^{*2}$. Hereinafter, the subscript e indicates parameters at the upper boundary of the boundary layer, and the asterisk stands for dimensional quantities. The acoustic conductivity A_y is expressed as [5]

$$A_y = -\frac{\phi}{Z_0}\tanh(\Lambda h), \quad (2)$$

where ϕ is porosity, h is the porous-layer thickness divided by the boundary-layer scale $l = \sqrt{\frac{\nu_e^* L^*}{U_e^*}}$, ν_e^* is the kinematic viscosity of the gas, and L^* is the distance from the leading edge of the model. The characteristic impedance Z_0 and propagation constant Λ are expressed in terms of the dimensionless dynamic density $\tilde{\rho} = \frac{\bar{\rho}^*}{\rho_w^*}$ and the dynamic compressibility $\tilde{C} = \gamma P_w^* \bar{C}^*$ as

$$Z_0 = \frac{\sqrt{\tilde{\rho}/\tilde{C}}}{M_e\sqrt{T_w}}, \quad \Lambda = \frac{i\omega M_e}{\sqrt{T_w}}\sqrt{\tilde{\rho}\tilde{C}}, \quad (3)$$

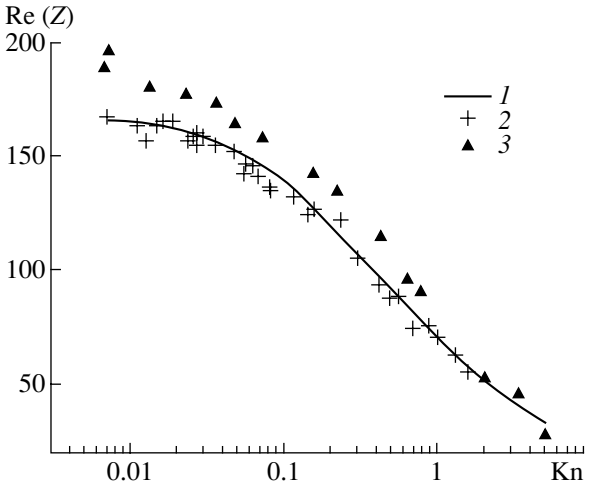


Fig. 3. Real part of the characteristic impedance $Z = \sqrt{\tilde{\rho}/\tilde{C}}$ (line) calculated as a function of the Knudsen number for $|\varsigma| = 10^{-2}$ in comparison with the values calculated with the measurement data [14] for (crosses) copper and (triangles) iron tubes.

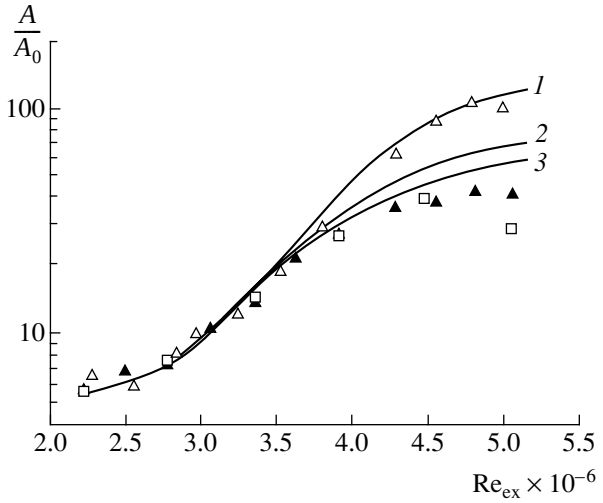


Fig. 4. Disturbances of the mass rate for the second mode on the smooth and porous walls for f = 275 kHz. Experimental data for (solid triangles) natural perturbations on the porous wall, (squares) artificial perturbations on the porous wall, and (open triangles) natural perturbations on the smooth wall. Lines *1–3* are the respective calculations.

where γ is the adiabatic index, the subscript w indicates quantities on the pore surface, and M_e is the Mach number on the boundary of the boundary layer.

For a continuum (where the mean free path in the gas filling a pore is much less than the pore radius), the characteristics of the porous medium $\tilde{\rho}$ and $\tilde{C}$ are obtained from the classical solution [11] of the problem of the propagation of perturbations in an individual cylindrical pore [12]. However, in the experiments under consideration, the radius of the pores is so small that the effects of gas rarefaction are significant. These effects were analyzed for small (but nonzero) Knudsen numbers characterizing the ratio of the mean free path to the pore radius. Molecular effects were assumed to be significant only in the relatively thin Knudsen layer on the pore surface. As in the classical case, perturbations of velocity and temperature in a pore are described by the linearized Navier–Stokes equations. However, boundary conditions on the walls of the pore are imposed with allowance for sliding effects and the temperature jump [13]

$$v_w^* = -B_v \mathrm{Kn}\left(\frac{\partial v^*}{\partial r}\right)_w, \quad \theta_w^* = -B_E \frac{\mathrm{Kn}}{\mathrm{Pr}}\left(\frac{\partial \theta^*}{\partial r}\right)_w. \quad (4)$$

Here, $\mathrm{Kn} = \dfrac{2\mu^*}{\rho_w^* \bar{c}^* r_0^*}$ is the Knudsen number, where $\bar{c}^* = \sqrt{\dfrac{8R_g T_w^*}{\pi}}$ is the average magnitude of the molecular velocity, R_g is the universal gas constant per unit mass, and T_w^* is the unperturbed temperature of the gas; Pr is the Prandtl number; $r = \dfrac{r^*}{r_0^*}$ is the radial coordinate divided by the pore radius; $B_v = 2\alpha_v^{-1} - 1$, where α_v is the accommodation coefficient for the tangential molecular momentum; and $B_E = \dfrac{2\gamma(2\alpha_E^{-1} - 1)}{\gamma + 1}$, where α_E is the energy accommodation coefficient. The pore-surface temperature is assumed to be constant.

Solving the above problem for perturbations of velocity and temperature, we obtain the following analytical expressions for the dynamic density and compressibility:

$$\tilde{\rho} = \frac{1}{1 - F(B_v, \varsigma)}, \quad \tilde{C} = 1 + (\gamma - 1)F(B_E, \varsigma\sqrt{\mathrm{Pr}}). \quad (5)$$

Here,

$$F(B_E, \varsigma) = \frac{Q(\varsigma)}{1 - 0.5B_v\varsigma^2 Q(\varsigma)},$$

$$F(B_E, \varsigma\sqrt{\mathrm{Pr}}) = \frac{Q(\varsigma\sqrt{\mathrm{Pr}})}{1 - 0.5B_E(\varsigma\sqrt{\mathrm{Pr}})^2 Q(\varsigma\sqrt{\mathrm{Pr}})},$$

$\varsigma = \sqrt{\dfrac{i\omega^* \rho_w^* r_0^{*2}}{\mu_w^*}}$, $Q(\varsigma) = \dfrac{2J_1(\varsigma)}{\varsigma J_0(\varsigma)}$, and $J_{0,1}(\varsigma)$ are the Bessel functions. Expressions (5) for Kn = 0 become the known expressions for a continuum (see, e.g., [12]).

To illustrate the applicability of the above expressions, Fig. 3 shows the real part of the characteristic impedance calculated as a function of the Knudsen number for the dimensionless parameter $|\varsigma| = 10^{-2}$ (low-frequency band, where gas-rarefaction effects are most

pronounced) in comparison with values calculated with the measurement data from [14]. Since the accommodation coefficients were not determined in that work, the calculations were performed with the values $\alpha_v = \alpha_E = 0.9$, which approximately correspond to the experimental values for the materials and temperature intervals indicated in [14]. Similar dependences were obtained for all other parameters used in the calculations. As is seen in Fig. 3, the theory agrees well with experiment over a much wider Knudsen-number range (up to $\mathrm{Kn} \approx 5$) than is assumed in the theoretical model under consideration.

Stability under two-dimensional second-mode perturbations ($\beta = 0$) was analyzed for the flow parameters corresponding to the experimental data. The gas-rarefaction effect is shown to increase the acoustic conductivity A_y of pores. In turn, this effect increases the efficiency of the UAC.

Figure 4 shows theoretical growth curves (lines) in comparison with experimental data (points). The experimental data correspond to the maxima in distributions similar to those shown in Fig. 2. The theoretical curves are normalized to the experimental points in the initial *X* section. Agreement with the experimental data is good on the smooth cone surface (cf. line *1* and open triangles). Growth curve *2* for the porous wall lies above the experimental points (closed triangles and squares). Under the assumption that this discrepancy is attributable to the conic shape of the pores, growth curve *3* was calculated for an average pore radius $r_0^* = 28.5$ μm. The theory with this correction agrees satisfactorily with the experimental data.

In summary, the theoretical analysis and measurements of stability on the sharp cone in the wind tunnel for $M = 6$ corroborated that the hypersonic boundary layer is stabilized by a passive porous coating with a regular microstructure. If the roughness of the surface decreases and the primarily two-dimensional flow regime is ensured (in order to avoid the instability of transverse flows), it is possible to stabilize the second mode and thereby to ensure the laminar flow regime over the most part of an aircraft. The theoretical and experimental results of this work can be used to estimate the characteristics of the UAC and to design a system for the passive laminarization of the hypersonic boundary layer.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 02-01-00141) and the European Office of Aerospace Research and Development with the International Science and Technology Center (project no. 2172).

REFERENCES

1. L. M. Mack, AGARD Rept., No. 709, 3 (1984).
2. V. I. Lysenko and A. A. Maslov, J. Fluid Mech. **147**, 39 (1984).
3. M. R. Malik, AIAA J. **27** (11), 1487 (1989).
4. N. D. Malmuth, A. V. Fedorov, V. Shalaev, *et al.*, AIAA Pap., No. 98-2695 (1998).
5. A. V. Fedorov, N. D. Malmuth, A. Rasheed, and H. G. Hornung, AIAA J. **39** (4), 605 (2001).
6. A. Rasheed, H. G. Hornung, A. V. Fedorov, and N. D. Malmuth, AIAA Pap., No. 2001-0274 (2001).
7. V. M. Fomin, A. V. Fedorov, A. N. Shiplyuk, *et al.*, Dokl. Akad. Nauk **384** (2), 197 (2002) [Dokl. Phys. **47**, 401 (2002)].
8. A. A. Maslov, A. N. Shiplyuk, A. A. Sidorenko, and D. Arnal, J. Fluid Mech. **426**, 73 (2001).
9. A. V. Fedorov and A. P. Khokhlov, Theor. Comput. Fluid Dyn. **15** (4), 231 (2002).
10. A. Fedorov, A. Shiplyuk, A. Maslov, *et al.*, J. Fluid Mech. **479**, 99 (2003).
11. G. Kirchoff, Poggendorfer Ann. **134**, 177 (1868).
12. C. Zwikker and C. W. Kosten, *Sound Absorbing Materials* (Elsevier, New York, 1949).
13. S. Chapmen and S. Cowling, *Mathematical Theory of Non-Uniform Cases*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1952; Inostrannaya Literatura, Moscow, 1960).
14. G. K. Brown, A. DiNardo, G. K. Cheng, and T. K. Sherwood, J. Appl. Phys. **17** (10), 802 (1946).

Translated by R. Tyapaev

Doklady Physics, Vol. 49, No. 12, 2004, pp. 768–771. Translated from Doklady Akademii Nauk, Vol. 399, No. 6, 2004, pp. 770–774.

MECHANICS

Griffith Formula for a Crack in a Piezoelectric Body

A. A. Kulikov* and S. A. Nazarov**

Presented by Academician N.F. Morozov May 27, 2004

Received June 15, 2004

ENERGY AND ENTHALPY FUNCTIONALS

We consider a plane homogeneous piezoelectric body Ω with a straight-line crack Λ that is assumed for definiteness to be a boundary crack. The origin of the Cartesian $x = (x_1, x_2)$ and polar (r, θ) coordinates is at the apex $\mathbb{O}$ of the crack, and the positive semiaxis $\mathbb{O}x_1$ and polar axis are directed along its continuation. For an arbitrary anisotropy of the material, we introduce elastic strain ε and stress σ tensors, an electric field strength **E**, and an electric displacement **D**, which are related as (see, e.g., [1, 2])

$$\begin{aligned}\sigma_{ij} &= a_{ij,\,pq}\varepsilon_{pq} - b_{ij,\,l}E_l,\\ D_k &= b_{pq,\,k}\varepsilon_{pq} + c_{k,\,l}E_l.\end{aligned} \tag{1}$$

The tensors a, b, and c of the physical constants of the material have the usual symmetric properties. In addition, the first two tensors are positive. Hereinafter, the summation over the repeated indices taking values 1, 2, and 3 is implied. However, since the problem is two-dimensional, all quantities are independent of the x_3 coordinate, whose axis is directed perpendicularly to the (x_1, x_2) plane. Volume forces are absent. An external load g and an electric displacement G are given at the part Σ of the boundary $\partial\Omega$, and they vanish at the faces $\Lambda_\pm$. The external displacements ξ and the electric potential Ξ are specified at the remaining part $\Gamma = \partial\Omega\backslash\bar{\Sigma}$, which is far from the crack. For the simplicity of calculations, we assume that Γ is a nonempty arc.

The potential energy $\mathcal{U}$ accumulated in the body is equal to $\mathscr{E} - \mathscr{A}$, where the free energy $\mathscr{E}$ and work $\mathscr{A}$ of external forces have the form

$$\mathscr{E} = \frac{1}{2}\int_\Omega (\sigma_{ij}\varepsilon_{ij} + D_kE_k)dx, \tag{2.1}$$

$$\mathscr{A} = \int_\Sigma (g_ku_k - G\varphi)ds - \int_\Gamma \nu_j(\sigma_{ij}\xi_i - D_j\Xi)ds. \tag{2.2}$$

Here, $\nu = (\nu_1, \nu_2, 0)$ is the outer normal unit vector, $\mathbf{u} = (u_1, u_2, u_3)$ is the displacement vector, and φ is the electric potential,

$$\varepsilon_{ij} := \varepsilon_{ij}(u) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right),\ E_k := E_k(\varphi) = -\frac{\partial\varphi}{\partial x_k}. \tag{3}$$

Substituting Eq. (3) into Eq. (1), we obtain a system of equilibrium equations and an electrostatic equation in the form

$$-\frac{\partial}{\partial x_j}a_{ij,\,pq}\frac{\partial u_p}{\partial x_q} - \frac{\partial}{\partial x_j}b_{ij,\,l}\frac{\partial\varphi}{\partial x_l} = 0, \quad i = 1, 2, 3, \quad \text{in} \quad \Omega, \tag{4.1}$$

$$-\frac{\partial}{\partial x_k}b_{pq,\,k}\frac{\partial u_p}{\partial x_q} + \frac{\partial}{\partial x_k}c_{k,\,l}\frac{\partial\varphi}{\partial x_l} = 0 \quad \text{in} \quad \Omega. \tag{4.2}$$

These equations, together with the boundary conditions

$$\nu_ja_{ij,\,pq}\frac{\partial u_p}{\partial x_q} + \nu_jb_{ij,\,l}\frac{\partial\varphi}{\partial x_l} = g_i, \quad i = 1, 2, 3, \ \text{on} \ \Sigma, \tag{5.1}$$

$$\nu_kb_{pq,\,k}\frac{\partial u_p}{\partial x_q} - \nu_kc_{k,\,l}\frac{\partial\varphi}{\partial x_l} = G \ \text{on} \ \Sigma, \tag{5.2}$$

$$u_i = \xi_i, \quad i = 1, 2, 3, \quad \varphi = \Xi \ \text{on} \ \Gamma \tag{5.3}$$

specify a formally self-adjoint boundary value problem that has a polynomial property and is thereby elliptic [3, Section 1]. The symmetric quadratic form

$$\mathbf{E}(u, \varphi;\ v, \psi) = \frac{1}{2}\int_\Omega \{\sigma_{ij}(u, \varphi)\varepsilon_{ij}(v) - D_k(u, \varphi)E_k(\psi)\}dx, \tag{6}$$

which corresponds to this problem and which is constructed by the standard rules, is not sign-definite due to

* *Faculty of Mathematics and Mechanics, St. Petersburg State University, Bibliotechnaya pl. 2, Staryĭ Peterhof, St. Petersburg, 198904 Russia*
e-mail: akulikov@ok.ru

** *Institute of Problems of Mechanical Engineering, Russian Academy of Sciences, Vasil'evskiĭ Ostrov, Bol'shoĭ pr. 61, St. Petersburg, 199178 Russia*
e-mail: serna@snark.ipme.ru

1028-3358/04/4912-0768$26.00 © 2004 MAIK "Nauka/Interperiodica"

the minus sign of the second term in the integrand. Nevertheless, changing unknowns (see example 1.13 in [3]) and applying the Lax–Milgram lemma, along with Korn's inequality, one finds that the problem specified by Eqs. (4) and (5) in the variational formulation

$$2\mathbf{E}(u, \varphi; v, \psi) = \int_\Sigma (g_k v_k + G\psi) ds \quad \forall \{v, \psi\} \in \mathring{H}^1(\Omega; \Gamma)^4 \tag{7}$$

with any right-hand side

$$\{g, G\} \in L_2(\Sigma)^4, \quad \{\xi, \Xi\} \in H^{1/2}(\Gamma)^4 \tag{8}$$

has the single solution $\{u, \varphi\}$ that satisfies the relation

$$\{u^0, \psi^0\} := \{u - \hat{\xi}, \varphi - \hat{\Xi}\} \in \mathring{H}^1(\Omega; \Gamma)^4$$

and the estimate

$$\|\{u, \varphi\}; H^1(\Omega)\| \le c\mathcal{N},$$

where $\mathcal{N}$ is the sum of the norms of functions (8) in the indicated spaces. Here, we use the usual notation for Sobolev–Slobodetskiĭ classes, ° indicates that homogeneous Dirichlet conditions (5.3) are satisfied, and $\hat{\xi}$ and $\hat{\Xi}$ are the continuations of fields ξ and Ξ, respectively, into the domains Ω in the class H^1.

The free energy $\mathcal{E}(u, \varphi)$ and enthalpy $\mathbf{E}(u, \varphi) = \mathbf{E}(u, \varphi; u, \varphi)$ are related as

$$\mathcal{E}(u, \varphi) = \mathbf{E}(u, \varphi) + \int_\Omega D_k(u, \varphi) E_k(\varphi) dx. \tag{9}$$

The difference

$$\mathbf{U}(u, \varphi) = \mathbf{E}(u, \varphi) - \mathbf{A}(u, \varphi) \tag{10}$$

is called the functional of the *mathematical* potential energy. Here,

$$\mathbf{A}(v, \psi) = \int_\Sigma (g_k v_k + G\psi) ds - \int_\Gamma v_j(\sigma_{ij}(v, \psi)\xi_i + D_j(v, \psi)\Xi) ds, \tag{11}$$

is the *mathematical* work of external forces, which differs from the *physical* energy $\mathcal{A}(v, \psi)$ (2.2). The latter integral in Eq. (11) is treated as $2\mathbf{E}(\hat{\xi}, \hat{\Xi}; v, \psi)$ for nonsmooth v or ψ. Wrong signs appear in Eqs. (6) and (11) because, according to the second definition in Eq. (3), Eq. (4.2) must be multiplied by $-\psi$ rather than ψ as was done for the formation of identity (7).

In this paragraph, $\xi = 0$ and $\Xi = 0$ (otherwise, the pair $\{u^0, \varphi^0\}$ must be used instead of $\{u, \varphi\}$). The solution $\{u, \varphi\} \in \mathring{H}^1(\Omega; \Gamma)^4$ of the problem specified by Eq. (7) or Eqs. (4) and (5) is the stationary point of the functional of *mathematical* potential energy (10). It is easy to see that this point is a *saddle point* and is not generally a stationary point of the functional of the *physical* potential energy. The latter has a minimum, but it is reached on the solution $\{u^\bullet, \varphi^\bullet\} \in \mathring{H}^1(\Omega; \Gamma)^4$ of the split problem: $b_{ij,k} = 0$ is set in Eqs. (4) and (5). In other words, for fixed elastic moduli $a_{ij,pq}$ and dielectric constants $c_{k,l}$, the body Ω has the minimum potential energy $\mathcal{U}$ if the piezoelectric constants $b_{ij,k}$ are equal to zero; i.e., elastic and electric fields do not interact.

According to Eqs. (6), (9)–(11), the functionals $\mathcal{U}$ and $\mathbf{U}$ calculated on the solution $\{u, \varphi\}$ of the problem specified by Eqs. (4) and (5) are related as

$$\mathcal{U}(u, \varphi) = \mathbf{U}(u, \varphi) + \mathbf{A}^E(u, \varphi) = \mathbf{U}(u, \varphi) - \mathcal{A}^E(u, \varphi), \tag{12}$$

where $\mathbf{A}^E(u, \varphi) = -\mathcal{A}^E(u, \varphi)$ is given by Eq. (2.2) for $g = 0$ and $\xi = 0$, i.e., the work of external electric forces.

INCREMENT OF THE MATHEMATICAL ENERGY WITH GROWTH OF THE CRACK

According to the general results given in [4], there are four linearly independent power solutions

$$X^n(x) = r^{1/2}\mathcal{X}^n(\theta), \quad n = 1, 2, 3, 4, \tag{13}$$

of the homogeneous model problem specified by Eqs. (4), (5.1), and (5.2) on the plane with the semi-infinite cut $\{x: x_1 \le 0, x_2 = 0\}$. They generate the root singularities of stresses and displacement, and other singularities are absent. According to [5], the power solutions in Eq. (13) can be normalized by the conditions

$$\sigma_{2k}(X^n; r, 0) = (2\pi r)^{-1/2}\delta_{k,n}, \quad D_2(X^n; r, 0) = (2\pi r)^{-1/2}\delta_{4,n}, \tag{14}$$

where $n = 1, 2, 3, 4$; $k = 1, 2, 3$; and $\delta_{k,n}$ is the Kronecker delta, and determine the intensity coefficients of the stresses and displacement

$$K_k = \lim_{r \to +0} (2\pi r)^{1/2}\sigma_{2k}(u, \varphi; r, 0),$$

$$K_4 = \lim_{r \to +0} (2\pi r)^{1/2} D_2(u, \varphi; r, 0).$$

Owing to the results presented in [6], basis (13) provides a unique basis

$$Y^m(x) = r^{-1/2}\mathcal{Y}^m(\theta), \quad m = 1, 2, 3, 4, \tag{15}$$

in the lineal of power solutions of the model problem where the exponents are equal to –1/2. The following biorthogonality conditions are valid:

$$\mathbf{Q}(X^n, Y^m; \Upsilon) = \delta_{n,m}, \quad n, m = 1, 2, 3, 4, \qquad (16)$$

where Υ is the simple arc that connects the surfaces of the cut and encloses the vertex $\mathcal{O}$ and where the invariant integral has the form

$$\mathbf{Q}(u, \varphi; v, \psi; \Upsilon) = \int_\Upsilon \nu_j \{\sigma_{ij}(u, \varphi) v_i + D_j(u, \varphi)\psi - \sigma_{ij}(v, \psi)u_i - D_j(v, \psi)\varphi\} ds.$$

By definition, the weight functions $Z^n = \{z^n, \zeta^n\}$ are non-energy solutions of the homogeneous problem given by Eqs. (4) and (5) with the following prescribed increase near the vertex $\mathcal{O}$:

$$Z^n(x) = Y^n(x) + O(1), \quad r \longrightarrow +0.$$

They appear in the intensity coefficients in the integral representations [6]

$$K_n = \int_\Sigma (g_k z_k^n + G\zeta^n) ds - \int_\Gamma \nu_j (\sigma_{ij}(Z^n)\xi_i + D_j(Z^n)\Xi) ds = \mathbf{A}(z^n, \zeta^n) \qquad (17)$$

following from normalization conditions (16).

The derivatives $\dfrac{\partial X^n}{\partial x_1}$ are also power solutions of the model problem, but acquire singularities $O(r^{-1/2})$ in the crack vertex and are expanded into basis (15):

$$\frac{\partial X^n}{\partial x_1}(x) = -\sum_{m=1}^{4} M_{nm} Y^m(x). \qquad (18)$$

According to [7] and taking Eqs. (16) and (18) into account, we obtain

$$M_{nm} = \mathbf{Q}(\partial_1 X^n, X^m; \Upsilon) = \mathbf{Q}(\partial_1 X^m, X^n; \Upsilon) = M_{nm}, \qquad (19)$$

i.e., the 4×4 matrix M is symmetric. In contrast to the pure elastic problem, it is impossible to prove the positive definiteness of the matrix M (cf. [5, 7]). Moreover, when the cross terms in Eq. (1) are small, three eigenvalues of the matrix M are positive and one eigenvalue is negative. We note that the number of power solutions (13) and the corresponding coefficients of the stress intensity decreases to three, and the corresponding 3×3 matrix M is positive definite when the surfaces are in electric contact.

Let $\{u^h, \varphi^h\}$ be a solution of the problem given by Eqs. (4) and (5) on the crack Λ^h growing by a small length $h > 0$. Applying the method of matched expansions by the scheme given in [5, 7, 8], we find that the far field (approximation beyond the vicinity of the vertex $\mathcal{O}$) has the form

$$\{u^h, \varphi^h\} = \{u, \varphi\} + h \sum_{m,n=1}^{4} K_m M_{mn} \{z^n, \zeta^n\} + \ldots. \qquad (20)$$

Here, an ellipsis stands for minor terms, but estimates of residuals, which require weight norms (cf. [9, 10]), are not performed for brevity. Then, using Eqs. (11), (17), and (20), we calculate the increment of functional (10):

$$\Delta\mathbf{U} = \mathbf{U}^h - \mathbf{U} = -\frac{1}{2}(\mathbf{A}(u^h, \varphi^h) - \mathbf{A}(u, \varphi)) = -\frac{h}{2}\sum_{m,n=1}^{4} K_m M_{mn} \mathbf{A}(z^n, \zeta^n) + O(h^2) = -\frac{h}{2}\sum_{m,n=1}^{4} K_m M_{mn} K_n + O(h^2), \qquad (21)$$

where the estimate of the residual is assured by the papers mentioned.

INCREMENT OF THE PHYSICAL POTENTIAL ENERGY

Using Eqs. (12) and (21) and calculating the derivative $\dfrac{\partial \varphi^h}{\partial h}$ for $h = 0$ by using representation (20) (cf. [8]), we arrive at the relation

$$\Delta\mathcal{U} = \Delta\mathbf{U} + \mathbf{A}^E(u^h - u, \varphi^h - \varphi) = -\frac{h}{2}\sum_{m,n=1}^{4} K_m M_{mn}(K_n - 2N_n) + O(h^2) = -\frac{h}{2}\sum_{m,n=1}^{4} \{K_m^M M_{mn} K_n^M - K_m^E M_{mn} K_n^E\} + O(h^2). \qquad (22)$$

Here, K_m^M and K_m^E are the intensity coefficients generated by the elastic and electric forces, respectively ($G = \Xi = 0$ and $g = \xi = 0$ are step-by-step substituted into Eq. (17)), and

$$N_n = \int_\Sigma G\zeta^n ds - \int_\Gamma \nu_j D_j(Z^n)\Xi ds \qquad (23)$$

is the *nonlocal* characteristic.

If $G = \Xi = 0$ in Eq. (5), Eqs. (21) and (22) are undistinguishable. However, Eq. (22) for $G \neq 0$ or $\Xi \neq 0$ differs strongly from the classical Griffith formula, because, in addition to the usual combination of the products of the intensity coefficients, it includes the nonlocal component

$$\mathbf{A}^E(\partial_h u^h|_{h=0}, \partial_h \varphi^h|_{h=0}),$$

which is treated as the increment of the work of external electric forces. The latter representation in Eq. (22) does not remove the mentioned nonlocality, because the decomposition $K_m = K_m^M + K_m^E$ is impossible in terms of fields near the mouth of the crack. These circumstances lead to serious consequences. First, applying the Griffith concept to piezoelectric bodies (cf. [1, Sect. 6]), we find the crack-growth condition

$$\sum_{m,n=1}^{4} K_m^M M_{mn} K_n^M - \sum_{m,n=1}^{4} K_m^E M_{mn} K_n^E = 4\gamma, \qquad (24)$$

where γ is the surface-energy density. The left-hand side of Eq. (24) is not sign-definite and can be made either positive or negative by the correlated variation of mechanical and electric forces (control over the quasistatic breakdown process) with constant total intensity coefficients K_m. Thus, in contrast to the pure elastic problem, the Griffith energy criterion for the piezoelectric medium *is not equivalent* to the generalized Irvin and Novozhilov criterion. For the same cause (the presence of a nonlocal component), the Eshelby–Cherepanov–Rice integral in the piezoelectric medium (see [1, Sect. 33]) *does not provide* the energy-release rate.

The above conclusions and result (24) contradict formulas from [1, Chapter 6]. However, careful analysis revealed the following errors in [1, p. 296]: an erroneous definition of work (an extra factor of 1/2) and an incorrect integration by parts in the reciprocity relation. The elimination of the above defects returns the integral $\mathbf{A}^E(u^h - u, \varphi^h - \varphi)$ to Eq. (33.23) in [1].

Owing to the general results given in [7], the expression of the energy increment in terms of the intensity coefficients, which are the characteristics of the stressed state in the mouth of the crack, is a prerogative of self-adjoint problems. In particular, the transition to the non-self-adjoint realization of the piezoelectric problem changes the key integral representations in Eq. (17), which, according to [6], include singular solutions of the adjoint problem and finally form the abnormal term $\mathbf{A}^E(u^h - u, \varphi^h - \varphi)$ in Eq. (22). If the external force is simple and depends linearly on no more than four parameters τ_j, a good choice of these parameters provides quantities (23) in the form of combinations of the intensity coefficients K_n, and the difference in the braces in Eq. (22) takes on the form $K_m \mathcal{M}_{mn}(\tau_j) K_n$. However, the coefficients $\mathcal{M}_{mn}(\tau_j)$ cannot be treated as the material constants, as was said about Eq. (34.48) in [1, p. 312]. At the same time, elements (19) of the matrix M are the *material constants*, which are easily reconstructed from basis (13) of the power solutions that satisfies normalization conditions (14).

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 03-01-00835) and the Lyapunov French–Russian Institute for Applied Mathematics and Computer Science (project no. 00-01).

REFERENCES

1. V. Z. Parton and B. A. Kudryavtsev, *Electroelasticity of Piezoelectric and Conductive Bodies* (Nauka, Moscow, 1988) [in Russian].
2. V. T. Grinchenko, A. F. Ulitko, and N. A. Shul'ga, *Mechanics of Coupled Fields in Construction Elements* (Naukova Dumka, Kiev, 1989) [in Russian].
3. S. A. Nazarov, Usp. Mat. Nauk **54** (5), 77 (1999).
4. M. Costabel and M. Dauge, Math. Nachr. **235**, 29 (2002).
5. S. A. Nazarov, Prikl. Mat. Mekh. **62** (3), 489 (1998).
6. V. G. Maz'ya and B. A. Plamenevskiĭ, Math. Nachr. **76**, 29 (1977).
7. S. A. Nazarov, in *Weight Functions and Invariant Integrals. Computational Mechanics of Deformed Solids* (Moscow, 1990), Vol. 1, pp. 17–31 [in Russian].
8. S. A. Nazarov, Mekh. Tverd. Tela, No. 3, 124 (1988).
9. S. A. Nazarov and O. R. Polyakova, Tr. MMO **57**, 16 (1996).
10. S. A. Nazarov and B. A. Plamenevskiĭ, *Elliptic Problems in Domains with a Piecewise Smooth Boundary* (Nauka, Moscow, 1991) [in Russian].

Translated by R. Tyapaev

Doklady Physics, Vol. 49, No. 12, 2004, pp. 772–777. Translated from Doklady Akademii Nauk, Vol. 399, No. 6, 2004, pp. 775–779.

MECHANICS

Instability of a Vortex Wake behind Wind Turbines

V. L. Okulov* and J. N. Sorensen**

Presented by Academician V.V. Kozlov May 20, 2004

Received May 27, 2004

Modern wind-power generators (wind turbines) often are grouped together in order to ensure an increased operating efficiency at the expense of using a large number of turbines in the same terrain area. However, in this case, one or several turbines can turn out to be in the wake beyond other turbines. As is shown in Fig. 1, the airflow beyond a wind turbine is represented by a system of intense rotational helical vortices determining the dynamics of a far wake. The wake produces a significant periodic load for the structure of a wind turbine present in the wake. This fact, naturally, results in a reduction of the turbine in-service time. However, as is well known, for certain operating regimes of wind turbines, the vortex wake becomes unstable and breaks down, as is shown in Fig. 1b. If a turbine is located in a stable vortex wake (Fig. 1a), then the interaction with it may have significant consequences than in the case of wake breakdown (Fig. 1b). Therefore, analysis of stability conditions for a vortex system simulating the vortex wake beyond wind turbines, screws, and aerodynamic propellers is of great practical importance.

At a sufficiently large distance behind a wind turbine or an aerodynamic propeller, the vorticity is concentrated in N blade-tip vortices of helical shape. The vortices are located on a cylindrical surface of radius R and have the same azimuth shift by the angle $\frac{2\pi}{N}$. Let each of the vortices have circulation Γ and a radius of the vortex core equal to ε. In addition, let a rectilinear vortex with the opposite total circulation $-N\Gamma$ exist along the system axis. Such an $(N + 1)$-vortex system with the constant helical pitch $2\pi l$ (or, in the dimensionless form, $\tau = l/R$) is the simplest model of a wake behind aerodynamic propellers and wind turbines [1]. In this approximation, the problem under consideration is reduced to the analysis of stability of the $(N + 1)$-vortex system that moves in uniform airflow at a constant wind speed V (Fig. 2). Since the flow behind the wind turbine has a wake-like profile, we analyze a system of left-handed helical vortices [2]. It is worth mentioning that the system of $(N + 1)$ vortices was studied previously only in the case of point vortices or rectilinear vortex filaments (the ultimate case of helical vortices with an infinitely large pitch) for which their instability was determined (see, e.g., [3]). It is clear that this result obtained for the simplest particular case ($\tau = \infty$) is inconsistent with the visualizations of wakes behind aerodynamic propellers and wind turbines (see, e.g., [4]). In addition, the linear analysis of stability for a simpler equilibrium configuration of N helical vortices without a central vortex, which model a pair, triplet, etc., of helical vortices arising in the tornado core after the vortex breakdown [5, 6] cannot be applied for solving the problem posed. Therefore, in the present paper, the stability analysis for wakes behind aerodynamic propellers, screws, and wind turbines is generalized to the case of a system consisting of $N + 1$ left-handed vortices.

In accordance with the ideas of [7], outside the vortex cores, the components of the velocity induced by the system of $N + 1$ left-handed vortices, which moves at constant wind speed V and additionally rotates in the field generated by the central circulation vortex $-N\Gamma$, can be written out in the form

$$u_r = \frac{\Gamma a}{\pi l^2}\sum_{n=1}^{N}\mathrm{Im}\left\{\begin{array}{l} H_1^{1,1}\left(\frac{r}{l}, \frac{R}{l}, \chi - \frac{2\pi n}{N}\right) \\ H_1^{1,1}\left(\frac{R}{l}, \frac{r}{l}, \chi - \frac{2\pi n}{N}\right)\end{array}\right\},$$

* *Kutateladze Institute of Thermophysics, Siberian Division, Russian Academy of Sciences, ul. Lavrent'eva 1, Novosibirsk, 630090 Russia*
e-mail: vokulov@mail.ru

** *Technical University of Denmark, DK-2800 Kongens Lyungby, Denmark*
e-mail: jns@mek.dtu.dk

1028-3358/04/4912-0772$26.00 © 2004 MAIK "Nauka/Interperiodica"

$$u_\theta = -\frac{N\Gamma}{2\pi r} + \sum_{n=1}^{N}\left(\frac{\Gamma}{2\pi r}\begin{Bmatrix}0\\1\end{Bmatrix} + \frac{\Gamma R}{\pi r l}\mathrm{Re}\begin{Bmatrix}H_1^{0,1}\left(\frac{r}{l},\frac{R}{l},\chi-\frac{2\pi n}{N}\right)\\ H_1^{1,0}\left(\frac{R}{l},\frac{r}{l},\chi-\frac{2\pi n}{N}\right)\end{Bmatrix}\right), \quad (1)$$

$$u_z = V - \sum_{n=1}^{N}\left(\frac{\Gamma}{2\pi l}\begin{Bmatrix}1\\0\end{Bmatrix} - \frac{\Gamma R}{\pi l^2}\mathrm{Re}\begin{Bmatrix}H_1^{0,1}\left(\frac{r}{l},\frac{R}{l},\chi-\frac{2\pi n}{N}\right)\\ H_1^{1,0}\left(\frac{R}{l},\frac{r}{l},\chi-\frac{2\pi n}{N}\right)\end{Bmatrix}\right).$$

Here, $H_M^{I,J}(x, y, \chi) = \sum_{n=1}^{\infty} m^M I_m^{\langle I\rangle}(mx) K_m^{\langle J\rangle}(my) e^{im\chi}$ for $x \le y$, $I_m^{(0)}(mx)$ and $K_m^{(0)}(my)$ are the modified Bessel functions, and $I_m^{(1)}(mx)$ and $K_m^{(1)}(my)$ are their derivatives. In (1), the upper and lower lines in braces correspond to $r < R$ and $r \ge R$, respectively. The calculation of trigonometric Kapteyn's series $H_M^{I,J}$ turns out to be associated with certain difficulties, especially as $r \longrightarrow R$. Therefore, we use a method of singularity separation [5, 6],

$$H_M^{I,J}(x, y, \chi) = S_M^{I,J}(x, y, \chi) + R_M^{I,J}(x, y, \chi), \quad (2)$$

where the principal part $S_M^{I,J}$ is determined as a sum of simple functions

$$S_M^{I,J} = \lambda^{I,J}\left[b_{M,0}^{I,J}\frac{e^{\xi+i\chi}}{(e^\xi - e^{i\chi})^2} + b_{M,1}^{I,J}\frac{e^{i\chi}}{e^\xi - e^{i\chi}} + b_{M,2}^{I,J}\ln(1-e^{-\xi+i\chi}) + b_{M,3}^{I,J}\mathrm{Li}_2(e^{-\xi+i\chi}) + b_{M,4}^{I,J}\mathrm{Li}_3(e^{-\xi+i\chi}) + b_{M,5}^{I,J}\mathrm{Li}_4(e^{-\xi+i\chi})\right], \quad (3)$$

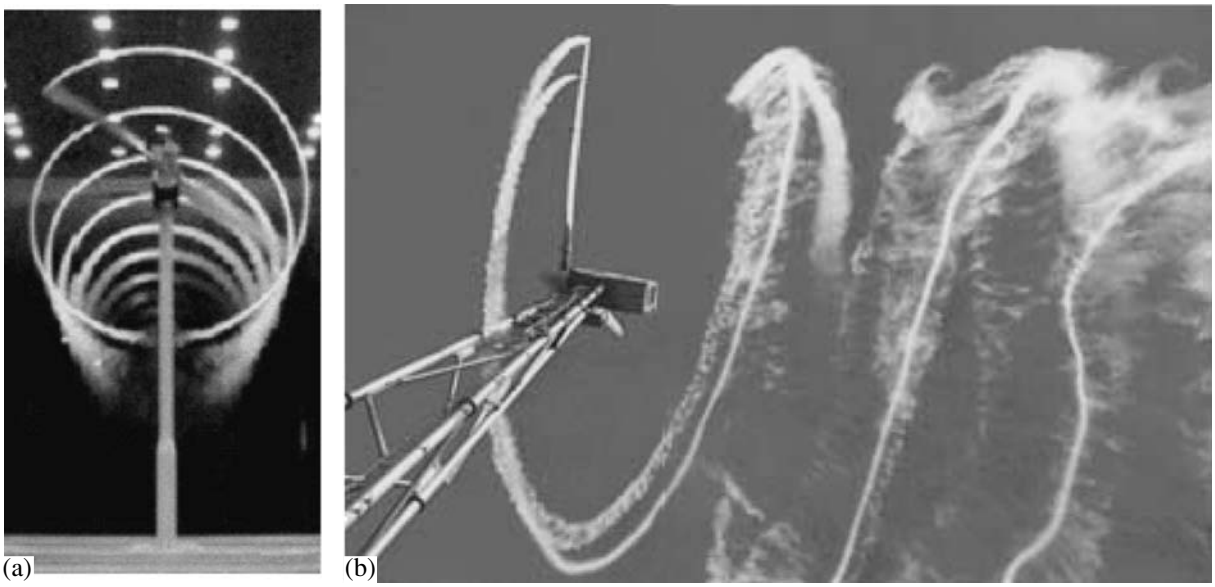
and $\mathrm{Li}_k(z) = \sum_{m=1}^{\infty}\frac{z^m}{m^k}$ ($|z| < 1$) are polylogarithms,

$$e^\xi = \frac{x}{y}\frac{\exp(\sqrt{1+x^2})(1+\sqrt{1+y^2})}{\exp(\sqrt{1+y^2})(1+\sqrt{1+x^2})};$$

$$\lambda^{I,J} = \frac{1}{2}\frac{(\sqrt{1+x^2})^{I-1/2}(\sqrt{1+y^2})^{J-1/2}}{x^I(-y)^J};$$

$$b^{I,J} = \begin{bmatrix} 0 & 0 & 1 & \alpha^{I,J} & \beta^{I,J} & \gamma^{I,J}\\ 0 & 1 & \alpha^{I,J} & \beta^{I,J} & \gamma^{I,J} & 0\\ 1 & \alpha^{I,J} & \beta^{I,J} & \gamma^{I,J} & 0 & 0\end{bmatrix}.$$

Furthermore, $\alpha^{I,J}$, $\beta^{I,J}$, and $\gamma^{I,J}$ are polynomials in terms of the rational functions ϑ_i and υ_i. These are obtained by multiplying asymptotic expansions (9.3.9), (9.3.10),



(a) (b)

Fig. 1. Visualization of a far wake behind a wind turbine: (a) stable regime [4] and (b) unstable regime accompanied by wake instability (RISO data, Denmark).

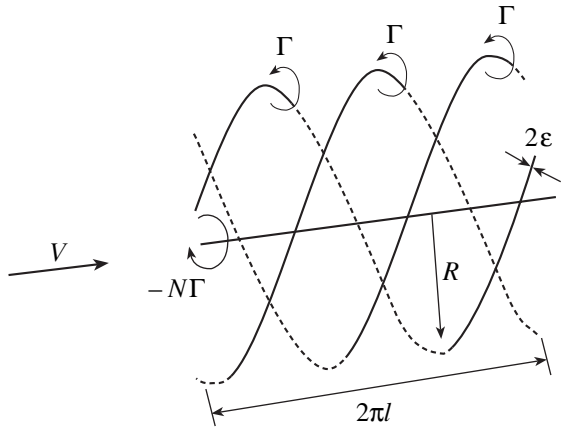


Fig. 2. Schematic diagram of an (N + 1) helical-vortex system.

(9.3.13), and (9.3.14) in [8]. For points taken in a circumference of radius $R\left(x=y=\frac{1}{\tau}\right)$, these polynomials have a simple form, namely,

$$\alpha^{\langle 0,1\rangle} = -\alpha^{\langle 1,0\rangle} = \frac{1}{2}\frac{\tau}{(\sqrt{1+\tau^2})^3}, \quad \alpha^{\langle 0,0\rangle} = \alpha^{\langle 1,1\rangle} = 0,$$

$$\beta^{\langle 0,0\rangle} = -\beta^{\langle 1,1\rangle} = \frac{1}{2}\frac{-\tau^2}{(1+\tau^2)^3}\left(\frac{\tau^2}{2}-\frac{1}{8}\right),$$

$$\beta^{\langle 0,1\rangle} = \beta^{\langle 1,0\rangle} = 0,$$

$$\gamma^{\langle 0,1\rangle} = -\gamma^{\langle 1,0\rangle} = \frac{1}{2}\frac{\tau^3}{(\sqrt{1+\tau^2})^9}\left(\tau^4-3\tau^2+\frac{3}{8}\right),$$

$$\gamma^{\langle 0,0\rangle} = \gamma^{\langle 1,1\rangle} = 0.$$

As a result, the singularities and their coefficients in two-dimensional representation (3) of solution (1) explicitly contain information on the torsion of helical vortices, whereas the regular residual $R_M^{I,J}$ in (2) can be ignored by virtue of its smallness [5, 6].

The unperturbed system of (N + 1) helical vortices participates in the translational motion and rotation about the cylinder axis with the total velocity u_b directed along the binormal to vortex filaments. This velocity is the sum of the self-induced velocity of each individual vortex (see Section 4.1.1 of [9]) and of the velocity induced by other vortices. By analogy with the problem of N helical vortices [5, 6], the desired velocity u_b can be written as

$$\frac{4\pi a}{\Gamma}u_b = \frac{N(1-\tau^2)}{\tau\sqrt{1+\tau^2}} + 1$$

$$+\frac{1}{1+\tau^2}\left[\ln\tau - \ln\left(N\varepsilon'\left(\sqrt{1+\tau^2}\right)^3\right) + \frac{3}{4}\right] + \frac{\tau}{(1+\tau^2)^4} \tag{4}$$

$$\times\left[\left(\tau^6-3\tau^4+\frac{3}{8}\tau^2\right)\frac{\zeta(3)}{N^2} - \tau^6 - 3\tau^4 - 3\tau^2 - 1\right]$$

$$-\left[\frac{2N\tau}{\sqrt{1+\tau^2}} - \frac{2b}{\sqrt{1+\tau^2}}\right],$$

where $\zeta(3) = 1.20206\ldots$ is the Riemann zeta-function, $b = \frac{V\cdot 2\pi R}{\Gamma}$ is the dimensionless wind speed, and ε' is the vortex-core radius normalized according to formula (3.14) of [9]. Equation (4), without the last term in square brackets, corresponds to the velocity of motion of a system containing N helical vortices and coincides with the results of [5, 6].

In order to analyze the stability of the system of (N + 1) helical vortices, we exploit an analogy with the analysis of linear stability for the equilibrium configuration of N vortices [5, 6]. Thus, we pass to the helical variables (r, χ) with the corresponding velocity components $\left(u_r, u_\chi = u_\theta + \frac{u_z}{\tau} \equiv -\frac{u_b\sqrt{1+\tau^2}}{\tau}\right)$. In other words, the given problem is reduced to the two-dimensional problem of the stability of the (N + 1)-vortex system with respect to infinitely small displacements of peripheral helical vortices from their equilibrium position at a fixed position of more intense central vortex. Let the kth vortex be displaced to the point $\left[R + r_k, \frac{2\pi k}{N} + t\left(\frac{u_\chi}{R}\right) + \chi_k\right]$. Then, in the linear approximation, the perturbed equations of motion for the kth vortex can be written as

$$\frac{dr_k}{dt} = \frac{\Gamma}{\pi R}\frac{1}{\tau^2}\mathrm{Re}\left\{\chi_k\sum_{n(n\neq k)} S_2^{1,1}\left(R, R, \frac{2\pi(n-k)}{N}\right)\right.$$

$$\left. - \sum_{n(n\neq k)} \chi_n S_2^{1,1}\left(R, R, \frac{2\pi(n-k)}{N}\right)\right\},$$

$$R\frac{d\chi_k}{dt} = \frac{\Gamma}{\pi R}\frac{1}{\tau^2} \times \mathrm{Re}\left\{ r_k \sum_{n(n\neq k)} \left[\frac{\tau^2+1}{\tau^2} S_2^{1,1}\left(R, R, \frac{2\pi(n-k)}{N}\right) + \frac{1-\tau^2}{\tau} S_1^{1,0}\left(R, R, \frac{2\pi(n-k)}{N}\right)\right] + \frac{(\tau^2+1)^2}{\tau^2} \times \sum_{n(n\neq k)} r_n S_2^{1,1}\left(R, R, \frac{2\pi(n-k)}{N}\right)\right\} + \frac{\sqrt{\tau^2+1}}{\tau} u_b r_k. \quad (5)$$

In formulas (5), the series $H_M^{I,J}$ entering into the first terms of the expansion of the velocity field were replaced by their principal parts, $S_M^{I,J}$, whereas the small residuals $R_M^{I,J}$ were ignored. For $0 \le k \le N-1$, following [10], we write the eigenfunctions of the dynamical system in the form

$$r_k = \alpha(t)\exp\frac{2\pi km}{N}, \quad \chi_k = \beta(t)\exp\frac{2\pi km}{N}, \quad (6)$$

where m is the subharmonic wave number taking integer values within the range $[0, N-1]$. For $m = 0$, the behavior of all vortices is the same. For other wave numbers, the substitution of (6) into (5) shows that solutions written out in the form (5) exist, provided that

$$\dot\alpha(t) = \beta(t)A(m), \quad \dot\beta(t) = \alpha(t)B(m). \quad (7)$$

Hence, it follows that α and β are proportional to $\exp(t\sqrt{AB})$ and the instability of the system depends on the sign of the product

$$\frac{16\pi^2R^3}{\Gamma^2}AB = \left[m(N-m)\frac{\sqrt{1+\tau^2}}{\tau} - \frac{\tau}{4}\frac{4\tau^2-3}{(\sqrt{1+\tau^2})^5} \times\left(\frac{N}{m} - C - \psi\left(-\frac{m}{N}\right)\right)\right]\left[m(N-m)\frac{(\sqrt{1+\tau^2})^3}{\tau^3} + 1 + \frac{\sqrt{1+\tau^2}}{\tau} - \frac{2b}{\tau} + \frac{1}{\tau\sqrt{1+\tau^2}}\left(\ln\frac{\tau}{N\varepsilon\sqrt{(1+\tau^2)}}\right) + \frac{3}{4} + \frac{\tau}{(\sqrt{1+\tau^2})^3}\left(\ln N - \left(1-\frac{1}{4\tau^2}\right)\left(\frac{N}{m} - C - \psi\left(-\frac{m}{N}\right)\right)\right)\right.$$
$$\left. + \frac{2N-3}{\tau^2} + \frac{\tau}{(\sqrt{1+\tau^2})^7}\left(\frac{\zeta(3)}{N^2}\left(\tau^4 - 3\tau^2 + \frac{3}{8}\right) - \tau^4 - 3\tau^2 - 3 - \frac{1}{\tau^2}\right) + \frac{\tau(1-\tau^2)}{(\sqrt{1+\tau^2})^9}\left(\tau^4 - 3\tau^2 + \frac{3}{8}\right)\frac{N^2-1}{N^2}\zeta(3)\right]. \quad (8)$$

Here, $C = 0.577215\ldots$ is the Euler constant. The psi-function $\psi(\cdot)$ for several arguments has the following values: $\psi(-1/2) = 0.03949\ldots$, $\psi(-1/3) = 1.68177\ldots$, $\psi(-2/3) = -1.63203\ldots$, $\psi(-1/4) = 2.91414\ldots$, and $\psi(-3/4) = -2.89412\ldots$. It is worth noting that, in the limiting case, as $\tau \to \infty$, product (8) has the form $m(N-m)[m(N-m)+2]$. This form entirely coincides with the result obtained in [3] for a system of point vortices, with the first multiplier of product (8) being always positive. Thus, if $B > 0$, for an arbitrary m within the range $0 < m \le N-1$, then the system is unstable, i.e., perturbations exponentially rise with time. If $B < 0$ for all m, then the system is linearly stable.

In order to apply this model for analysis of the far-wake stability of the wind turbine, we express geometric and kinematic wake parameters (τ, V, and Γ) in terms of the typical operating characteristics of a wind turbine (axial interference factor a of the wind speed and the tip speed ratio λ (i.e., the ratio between the velocity of blade-tip motion and the wind speed)). The parameter a in the wake can be determined as the ratio of the axial velocity induced by wake helical vortices and the wind speed. Ignoring the small diameter of vortex cores in the wake, we obtain, as a result of calculations of the mean axial velocity according to (1), $a = \frac{N\Gamma}{2\pi lV} = \frac{G}{\tau}$, where $G = \frac{N\Gamma}{2\pi RV}$ is the total dimensionless circulation of tip vortices. In the following simplification of the problem based on the lifting disk theory, we represent quantity G in the form $G = \frac{2a(1-a)}{\lambda}$. Thus, we find $\tau = \frac{2(1-a)}{\lambda}$. To replace wind speed V, we use its definition $\frac{V2\pi R}{\Gamma} = \frac{N}{G} = \frac{N\lambda}{2a(1-a)}$. As a result, the correlation equation expressed in terms of these parameters takes the form

$$\frac{1}{\vartheta(a,\lambda)}\left(2(1-a) + \frac{0.5\lambda^2}{1-a}\left(\ln\frac{2(1-a)}{N\varepsilon\vartheta(a,\lambda)} + \frac{7}{4}\right)\right) + \frac{(N-0.5)a - N}{2a(1-a)^2}\lambda^2 + \frac{2(1-a)\lambda^2}{\vartheta^3(a,\lambda)} \times\left(\ln N - \left(1 - \frac{\lambda^2}{16(1-a)}\right)\left(\frac{N}{m} - C - \psi\left(-\frac{m}{N}\right)\right)\right)$$

$$+ m(N-m)\frac{\vartheta^3(a,\lambda)}{8(1-a)^3} + \frac{(1-a)\lambda^6}{\vartheta^7(a,\lambda)}$$

$$\times\left(32\left(\frac{\zeta(3)}{N^2}-1\right)\frac{(1-a)^4}{\lambda^5} - 24\left(\frac{\zeta(3)}{N^2}+1\right)\frac{(1-a)^2}{\lambda^3}\right.$$

$$\left.+\left(\frac{3\zeta(3)}{4N^2}-6\right) - \frac{0.5\lambda^2}{(1-a)^2}\right) + \frac{(1-a)\lambda^6(4(1-a)^2-\lambda^2)}{\vartheta^9(a,\lambda)}$$

$$\times\left(32\frac{(1-a)^4}{\lambda^4} - 24\frac{(1-a)^2}{\lambda^2} + \frac{3}{4}\right)\frac{1-N^2}{N^2}\zeta(3) + 1 = 0,$$

where $\vartheta(a,\lambda) = \sqrt{4(1-a^2)+\lambda^2}$.

Figure 3 presents neutral stability curves for a system of $N + 1$ helical vortices modeling the far wake of wind turbines with two and three blades. The possible variations of controlling parameters for typical operating regimes of a wind turbine are denoted by shaded areas. As is seen, they lie, mainly, in the stable region. However, in some cases, wind turbines operate in conditions when the wake vortex structure becomes unstable, which is consistent with the observations in [11]. In the operating-regime zones, the effect of the vortex-core radius is weak and can be ignored.

We may conclude that the problem of stability for a system with $N + 1$ helical vortices is investigated in the analytical form. This system models the far wake behind the screws, aerodynamic propellers, and wind turbines. Typical operating regimes of a wind turbine were compared with characteristics determining the stability of the vortex system. As a result, we have found that, even with a rough estimate of the system vortex parameters in terms of turbine operating characteristics, the theory developed is qualitatively consistent with data from full-scale and model tests. The representation of the correlation equation in the simple analytical form can be considered a first step in finding recommendations for efficient control of wind-turbine clustered in parks, or "wind farms."

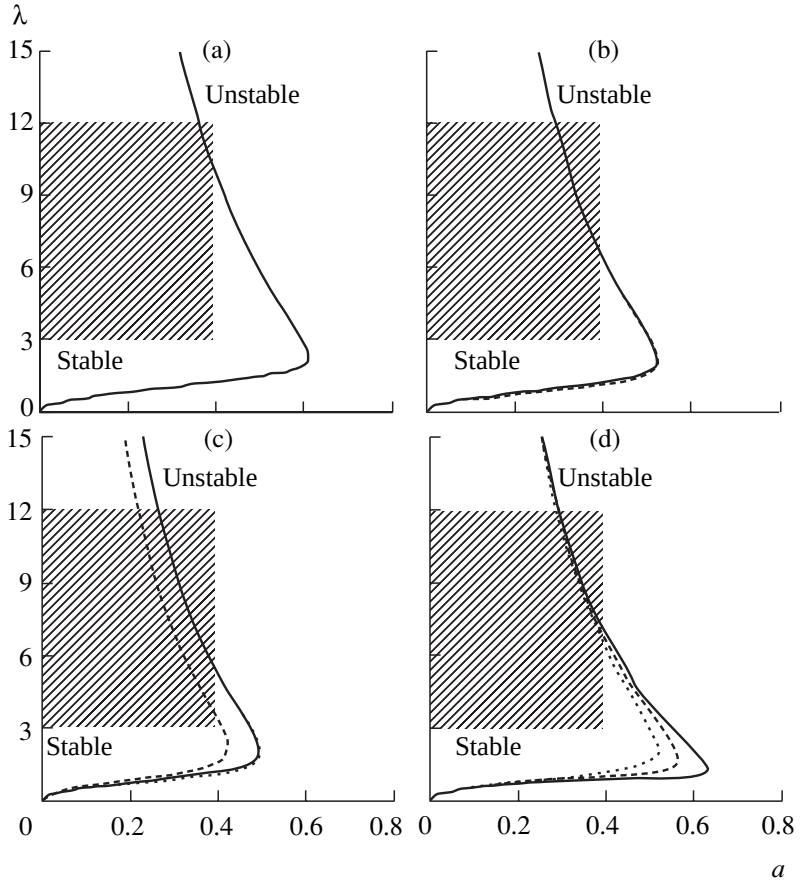


Fig. 3. Diagram of possible variations of wind-turbine operating parameters (shaded region) and curves of neutral stability for an $(N + 1)$ vortex system modeling the far wake behind the wind turbines with different numbers N of blades N and for different subharmonic wavenumbers m: (a) $N = 2$, $m = 1$ (solid line); (b) $N = 3$, $m = 1$ (solid line) and $m = 2$ (dashed line); (c) $N = 4$, $m = 1$ (solid line), $m = 2$ (dashed line), and $m = 3$ (dotted line); (d) effect of the vortex core radius for $N = 3$, $\varepsilon = 0.1$ (solid line), $\varepsilon = 0.01$ (dashed line), and $\varepsilon = 0.001$ (dotted line).

ACKNOWLEDGMENTS

The work was supported in part by the Russian Foundation for Basic Research, project no. 04-01-00124.

REFERENCES

1. N. E. Zhukovskiĭ, Trudy Otd. Fiz. Nauk Obshch. Lyubit. Estestvoznaniya **16** (1) (1912).
2. V. L. Okulov, Pis'ma Zh. Tekh. Fiz. **22** (19), 47 (1996).
3. G. K. Morikawa and E. V. Swenson, Phys. Fluids **14**, 1058 (1971).
4. L. J. Vermeer, J. N. Sorensen, and A. Crespo, Progress Aerospace Sci. **39**, 467 (2003).
5. V. L. Okulov, Pis'ma Zh. Tekh. Fiz. **28** (24), 80 (2002).
6. V. L. Okulov, J. Fluid Mech. **521**, 319 (2004).
7. J. C. Hardin, Phys. Fluids **25**, 1949 (1982).
8. M. Abramovitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, 1964).
9. R. L. Ricca, J. Fluid Mech. **273**, 241 (1994).
10. P. G. Saffman, *Vortex Dynamics* (Cambridge Univ. Press, Cambridge, 1992).
11. T. F. Pedersen and I. Antoniou, Wind Eng. **13**, 239 (1989).

Translated by G. Merzon