THE MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION

ST. PETERSBURG STATE POLYTECHNICAL UNIVERSITY

JOURNAL

Physics and Mathematics

VOLUME 16, No. 4, 2023

Peter the Great St. Petersburg
Polytechnic University
2023

ST. PETERSBURG STATE POLYTECHNICAL UNIVERSITY JOURNAL.
PHYSICS AND MATHEMATICS

The journal is included in the List of leading peer-reviewed scientific journals and other editions to publish major findings of theses for the research degrees of Doctor of Sciences and Candidate of Sciences.

The publications are presented in the VINITI RAS Abstract Journal and Ulrich's Periodical Directory International Database.

The journal is published since 2008 as part of the periodical edition 'Nauchno-tekhnicheskie vedomosti SPb-GPU'.

The journal is registered with the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR). Certificate ПИ № ФС77-52144 issued December 11, 2012.

The journal is distributed through the CIS countries catalogue, the «Press of Russia» joint catalogue and the «Press by subscription» Internet catalogue. The subscription index is **71823.**

The journal is in the **Web of Science** (Emerging Sources Citation Index), **Scopus**, the **Russian Science Citation Index** (RSCI) and the **Directory of Open Access Journals** (DOAJ) databases.

© Scientific Electronic Library (http://www.elibrary.ru).

Address: 195251 Politekhnicheskaya St. 29, St. Petersburg, Russia.

Phone: (812) 294-22-85.
http://ntv.spbstu.ru/physics

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

НАУЧНО-ТЕХНИЧЕСКИЕ ВЕДОМОСТИ

САНКТ-ПЕТЕРБУРГСКОГО ГОСУДАРСТВЕННОГО ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

Физико-математические науки

ТОМ 16, № 4

2023

Санкт-Петербургский политехнический
университет Петра Великого
2022

НАУЧНО-ТЕХНИЧЕСКИЕ ВЕДОМОСТИ САНКТ-ПЕТЕРБУРГСКОГО ГОСУДАРСТВЕННОГО ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА. ФИЗИКО-МАТЕМАТИЧЕСКИЕ НАУКИ

Журнал с 2002 г. входит в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученых степеней доктора и кандидата наук.

Сведения о публикациях представлены в Реферативном журнале ВИНИТИ РАН, в международной справочной системе «Ulrich's Periodical Directory».

С 2008 года выпускается в составе сериального периодического издания «Научно-технические ведомости СПбГПУ».

Журнал зарегистрирован Федеральной службой по надзору в сфере информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство о регистрации ПИ № ФС77-52144 от 11 декабря 2012 г.

Распространяется по Каталогу стран СНГ, Объединенному каталогу «Пресса России» и по Интернет-каталогу «Пресса по подписке». Подписной индекс **71823**.

Журнал индексируется в базе данных **Web of Science** (Emerging Sources Citation Index), **Scopus**, а также включен в базы данных **«Российский индекс научного цитирования»** (РИНЦ), размещенную на платформе Научной электронной библиотеки на сайте http://www.elibrary.ru, и **«Directory of Open Access Journals»** (DOAJ)

Адрес редакции и издательства:
Россия, 195251, Санкт-Петербург, ул. Политехническая, д. 29.
Тел. редакции (812) 294-22-85.
http://ntv.spbstu.ru/physics

SIMULATION OF PHYSICAL PROCESSES

Original article
DOI: https://doi.org/10.18721/JPM.17201

THE TESTING OF SOME EXTRA ACCURACY SCHEMES ON THE PROBLEM OF THE SHOCK WAVE – VORTEX INTERACTION

E. V. Babich✉, *E. V. Kolesnik*

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ lll.helen.lll@mail.ru

Abstract. In the paper, the calculation results of the shock wave – isoentropic vortex interaction have been presented. A complex nonstationary shock wave pattern was analyzed. The influence of the scheme for approximating the convective flows on the solution accuracy was shown to be weak. The results of calculations conducted using some extra accuracy schemes are presented: the data obtained by the second-order TVD scheme and the fifth-order WENO one being compared. The most accurate reconstruction of the vortex structure after its interaction with the shock wave was found to be achieved in the calculations when taking the WENO scheme.

Keywords: shock wave interaction, isoentropic vortex, numerical simulation, finite volume method, extra accuracy scheme

Citation: Babich E. V., Kolesnik E. V., The testing of some extra accuracy schemes on the problem of the shock wave – vortex interaction, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 8–20. DOI: https://doi.org/10.18721/JPM.17201

Научная статья
УДК 519.6:533.6.011
DOI: https://doi.org/10.18721/JPM.17201

ТЕСТИРОВАНИЕ СХЕМ ПОВЫШЕННОЙ ТОЧНОСТИ НА ЗАДАЧЕ О ВЗАИМОДЕЙСТВИИ УДАРНОЙ ВОЛНЫ С ВИХРЕМ

Е. В. Бабич✉, *Е. В. Колесник*

Санкт-Петербургский политехнический университет Петра Великого,

Санкт-Петербург, Россия

✉ lll.helen.lll@mail.ru

Аннотация. Представлены результаты решения задачи о взаимодействии ударной волны и изоэнтропического вихря. Проанализирована сложная нестационарная ударно-волновая картина. Показано слабое влияние схемы аппроксимации конвективных потоков на точность расчетов. Представлены результаты расчетов, полученных с применением схем повышенной точности: проведено сравнение результатов, полученных с использованием схем TVD второго порядка и WENO пятого порядка. Установлено, что наиболее точное восстановление структуры вихря после его взаимодействия с ударной волной достигается при расчетах с использованием схемы WENO.

Ключевые слова: ударно-волновое взаимодействие, изоэнтропический вихрь, численное моделирование, метод конечных объемов, схема повышенной точности

Ссылка для цитирования: Бабич Е. В., Колесник Е. В. Тестирование схем повышенной точности на задаче о взаимодействии ударной волны с вихрем // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 2. С. 8–20. DOI: https://doi.org/10.18721/JPM.17201

Introduction

One of the central problems of modern gas dynamics concerns the initiation and evolution of turbulent supersonic flows, arising, in particular, for vortex shedding from supersonic airfoil [1], for intake vortices shed by upstream surfaces [1], or for two intersecting bow shocks producing thin vortex layers that interact with reflected shock waves.

Problems of this type are generally very complex, and modern approaches are needed to accurately resolve various gas dynamic structures. For instance, simulation methods should provide such numerical dissipation which, on the one hand, would allow suppressing non-physical oscillations near discontinuities, and on the other hand, would be sufficiently small so as not to interfere with resolution of vortex structures.

The model problem of the interaction between shock waves and vortices is often used to test such methods [2]. When a plane shock wave collides with a vortex, this generates a perturbation propagating along the shock wave and deforming it. The flow field locally contracts and expands behind the curved shock wave, forming acoustic waves. These interesting phenomena occurring in laminar flow are closely related to interaction of shock waves and turbulent vortices, which is one of the main sources of noise.

Therefore, the problem of interaction between a shock wave and a vortex is of fundamental importance, with potential implications for accuracy estimates of various numerical schemes, which can later be used to solve more complex practical problems.

The goal of this study is to analyze the influence of schemes for approximating convective fluxes on the resolution quality of the flow structure, additionally evaluating the effectiveness of variable reconstruction schemes with increased accuracy.

The first experimental studies on the interaction between a vortex shed in the wake of a model airfoil and a shock wave propagating in a shock tube were carried out by Hollingsworth and Richards [3], and later by Dosanjh and Weeks [4]. The vortex was found to contract into an elliptical shape after passing through the shock wave front, with the major axis of the ellipse approximately equal to the diameter of the starting circular vortex, and the ratio of the major to the minor axis approximately equal to the ratio of densities on different sides of the shock wave. In turn, Naumann and Hermanns [5] conducted an experimental study in a shock tube, establishing that the deformation of the vortex corresponds to either regular or Mach reflection depending on the vortex intensity.

Guichard et al. [6] numerically solved the problem on the interaction of a vortex and a shock wave using a compact 6th-order scheme. The shape of the vortex was modified, and a triple point was detected, along with an increase in vorticity after the interaction. We should note that even though a large amount of evidence has been accumulated, some aspects, such as the influence of the applied scheme on the accuracy of the solution and the dependence of the solution structure on the intensity of the vortex, are yet to be fully explored [7].

Particulars of problem statement

Consider the two-dimensional problem of the interaction of a single cylindrical isentropic vortex with a shock wave. In this case, unsteady gas flow containing discontinuity surfaces, i.e., bow shocks (in addition to the main shock wave), and contact surfaces is simulated. The computational domain (Fig. 1,*a*) is a rectangle with the dimensions $[-1, 1] \times [0, 1]$. The coordinate system $x0y$ is associated with a shock wave whose front at the initial time ($t = 0$) is stationary and located in the section $x_s = 0$, while the center of a two-dimensional isentropic vortex is located at a point with the coordinates $(x_v, y_v) = (-0.5, +0.5)$.

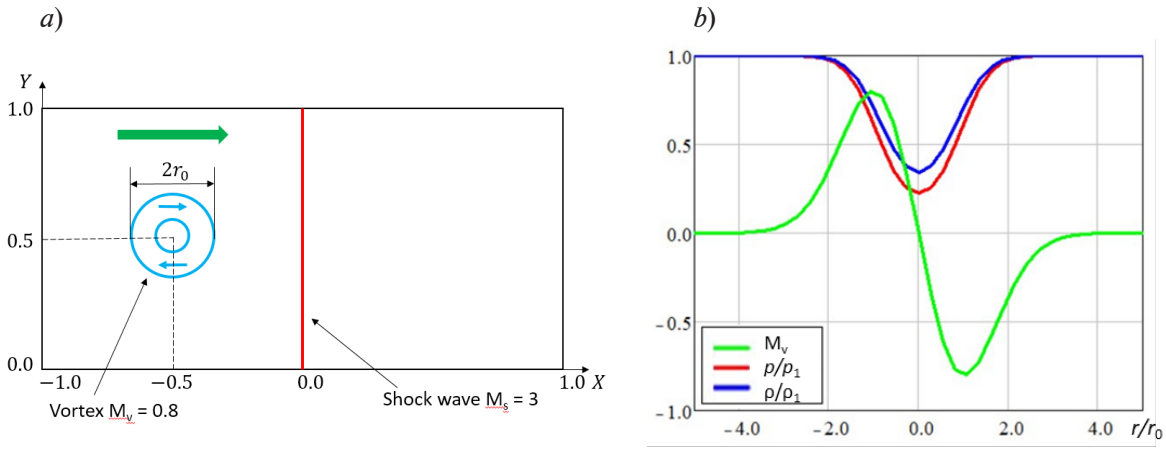

Fig. 1. Problem statement for interaction of shock wave with isentropic vortex: *a* corresponds to the computational domain (M_S is the shock-wave Mach number, r_0 is the effective radius of the vortex); *b* shows distributions of the vortex Mach number M_V, pressure and density (normalized quantities) in the vortex cross section
The green arrow shows the direction of the vortex propagation

The working fluid is a perfect gas with the adiabatic index $\gamma = 1.4$. The main parameters of the problem: M_S is the Mach number characterizing the intensity of the shock wave, $M_S = u_x/c$ (u_x is the flow velocity in front of the wave, c is the speed of sound); $V = V(r)$ is the velocity profile of the cylindrical vortex (r is the distance from the center of the vortex); M_V is the Mach number characterizing the intensity of the vortex, $M_V = V_m/c$ (V_m is the maximum velocity on the profile); r_0 is the effective radius of the vortex (determined by a point on the velocity profile where $V(r_0) = V_m$).

The velocity profile of the vortex (Fig. 1,*b*) is given by the formula

$$V(r) = V_m (r / r_0) \exp\left\{\frac{1-(r/r_0)^2}{2}\right\},$$

where $r = \sqrt{(x-x_V)^2 + (y-y_V)^2}$.

The remaining parameters of the problem have the following values: $M_S = 3$; $M_V = 0.8$; $r_0 = 0.075$.

The initial field is obtained by adding a cylindrical vortex to the background flow in front of the shock wave. First of all, let us define the shock wave field. The parameters before the shock wave front are subscripted by 1, and those behind the front are subscripted by 2:

$$u_{x1} = M_S \sqrt{\frac{\gamma p_1}{\rho_1}},\ u_{y1} = 0,\ u_{x2} = u_{x1} \frac{(\gamma-1)M_S^2 + 2}{(\gamma+1)M_S^2},\ u_{y2} = 0,$$

$$p_1 = 1,\ \rho_1 = 1,\ \rho_2 = \rho_1 \frac{(\gamma+1)M_S^2}{(\gamma-1)M_S^2 + 2},\ p_2 = p_1 \frac{2\gamma M_S^2 - (\gamma-1)}{(\gamma+1)},$$

where u are the velocities, p_i are the pressures, and ρ_i are the densities.

If a vortex is introduced, the parameters in the region in front of the shock wave are determined by the following formulas:

$$u_x = u_{x1} + (y - y_V) f / r_0,\ u_y = u_{y1} + (x - x_V) f / r_0,$$

$$\rho = \rho_1 g^{1/(\gamma-1)},\ p = p_1 g^{\gamma/(\gamma-1)},$$

where $f(r) = V_m \exp\left\{\frac{1-(r/r_0)^2}{2}\right\},\ g(r) = 1 - \frac{(\gamma-1)\rho_1}{2\gamma p_1}\cdot f^2.$

In accordance with the recommendations given in [2], the problem is solved taking into account the influence of gas viscosity, therefore allowing to eliminate the instability of contact discontinuities for high-order approximation schemes, contributing to a smoother flow structure. Respectively, the Reynolds and Prandtl numbers are equal: Re = 10^4, Pr = 3/4. The physical viscosity μ is assumed to be constant and equal to

$$\mu = \rho_1 u_{x1}(2r_0)/\mathrm{Re}.$$

A condition for flow symmetry is imposed at the upper and lower boundaries of the region, a supersonic inflow is set at the left boundary, and a subsonic outflow at a given pressure is set at the right boundary.

Computational aspects

The unstructured-grid finite-volume SINF/Flag-S code, developed at Peter the Great St. Petersburg Polytechnic University and validated in numerous tests over the years, was used for the computations [8].

The complete 2D Navier-Stokes equations for thermally and calorically perfect gas were solved. convective fluxes at the face of the computational cell in the Navier–Stokes equation were calculated by shock-capturing methods, smearing the discontinuity surfaces on a certain number of grid cells due to numerical viscosity (scheme dissipation) whose effect is similar to the effect of physical viscosity.

This study tested methods based on the exact or approximate solution of the Riemann problem on the breakdown of an arbitrary discontinuity (Godunov [9], Roe [10], HLL [11], HLLC [12]) as well as methods using flux vector splitting of the flow vector (in particular, the AUSM-family schemes [13]).

The reconstruction of variables at the face was carried out using quasi-one-dimensional second-order TVD schemes with a van Albada limiter [14]. Calculations were also carried out by the WENO5 scheme with fifth-order accuracy for smooth solutions on a uniform grid (proposed in [15]). The advantage of the WENO-family schemes, which are currently considered the most effective for problems with discontinuities, is that they combine increased accuracy for smooth solutions and limited generation of discontinuity oscillations [15].

The computational mesh consisted of square cells of size $h \times h$. A sequence of meshes with the step h = 1/200, 1/400, 1/500, 1/600, 1/800 was considered in this paper.

Computing resources of the Polytechnic Supercomputer Center (www.scc.spbstu.ru) were used for the computations.

Computational results

Structure of the flow. The evolution of the flow structure during the interaction of the vortex with the shock wave is shown in Fig. 2 (simulation by the AUSM scheme using a reference grid with a spacing h = 1/800, reconstruction was carried out by the second-order TVD scheme).

At the initial time, the vortex was located at some distance from the shock wave and rotated clockwise. At time t = 157.50 ms (Fig. 2,*b*), the vortex passed halfway through the shock wave, flattened due to high pressure and took the shape of an ellipse. As the pressure gradient inside the vortex is lower than in the external flow, the wave front should curve, since the pressure drop in the oblique shock wave is smaller.

By the time t = 183.75 ms (Fig. 2,*c*), the vortex completely passed through the shock wave front. A triple point appeared on the wave front; the Mach stem impinged on the vortex, separating the external region with elevated pressure from the region with reduced pressure inside the vortex. The intrinsic velocity of the vortex in its upper part is summed up with the velocity of incident flow, producing a region of supersonic flow (colored in black in Fig. 3,*c*).

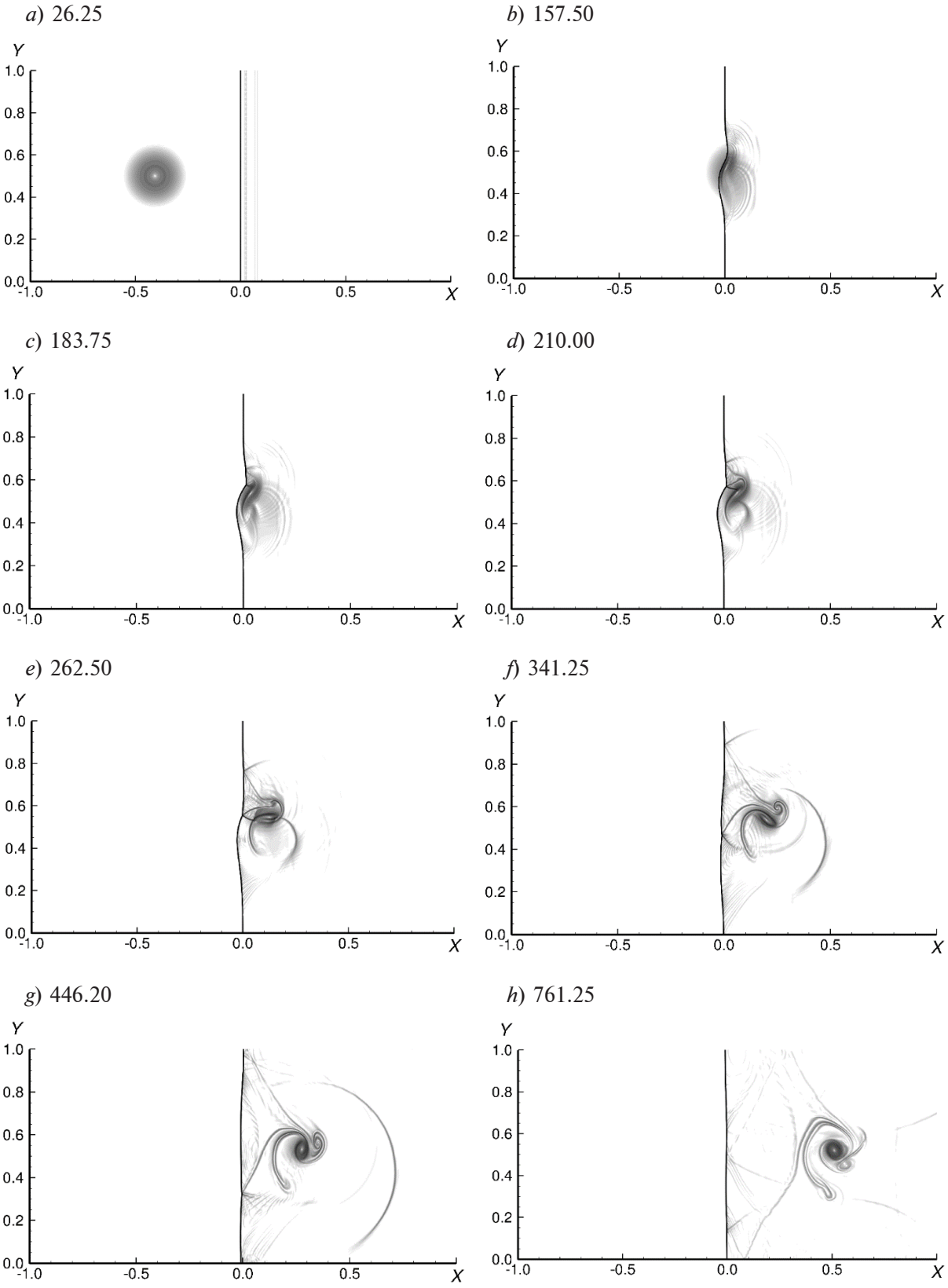

Fig. 2. Time evolution of problem solution: fields of the density gradient magnitude (Schlieren visualization). The time points (ms) are given in the graphs. The solution was obtained using the AUSM scheme and a grid with $h = 1/800$

After $t = 262.50$ ms (Fig. 2,*e*), the triple point generated two contact discontinuities: the first one continued to rotate clockwise (Fig. 2, *d*), and the second, weaker, one rotated in the opposite direction. The distorted vortex generated acoustic waves, some of which turned into weak shock waves with a curved front (Fig. 2,*f*). A configuration consisting of two contact discontinuities with opposite vorticity vectors rotated clockwise around some axis. As the vortex moved forward, its shape changed from elliptical to circular, but its intensity became twice as high as at the initial moment due to high pressure. As the vortex moved away, the shock wave front was smoothened, and the contact discontinuities separated from the triple point. The contact discontinuity rotating counterclockwise weakened significantly (Fig. 2,*h*).

The influence of grid spacing. Below we consider the influence of grid spacing h on the problem solution for the time when the vortex completely passed through the shock wave front ($t = 446.25$ ms). Fig. 3 shows the fields of the flow parameters at this time, calculated using the most refined of the grids considered ($h = 1/800$), and Fig. 4 shows schlieren images obtained by calculations using the AUSM scheme with a sequence of refined grids. Evidently, the two contact discontinuities are significantly further away from each other on the coarsest grid ($h = 1/200$) than in the reference solution for the grid with $h = 1/800$, while the resulting configuration is very close to the reference one for the grid with $h = 1/400$. Furthermore, the more refined the grid, the more pronounced the contact discontinuities, shock waves and weaker acoustic perturbations.

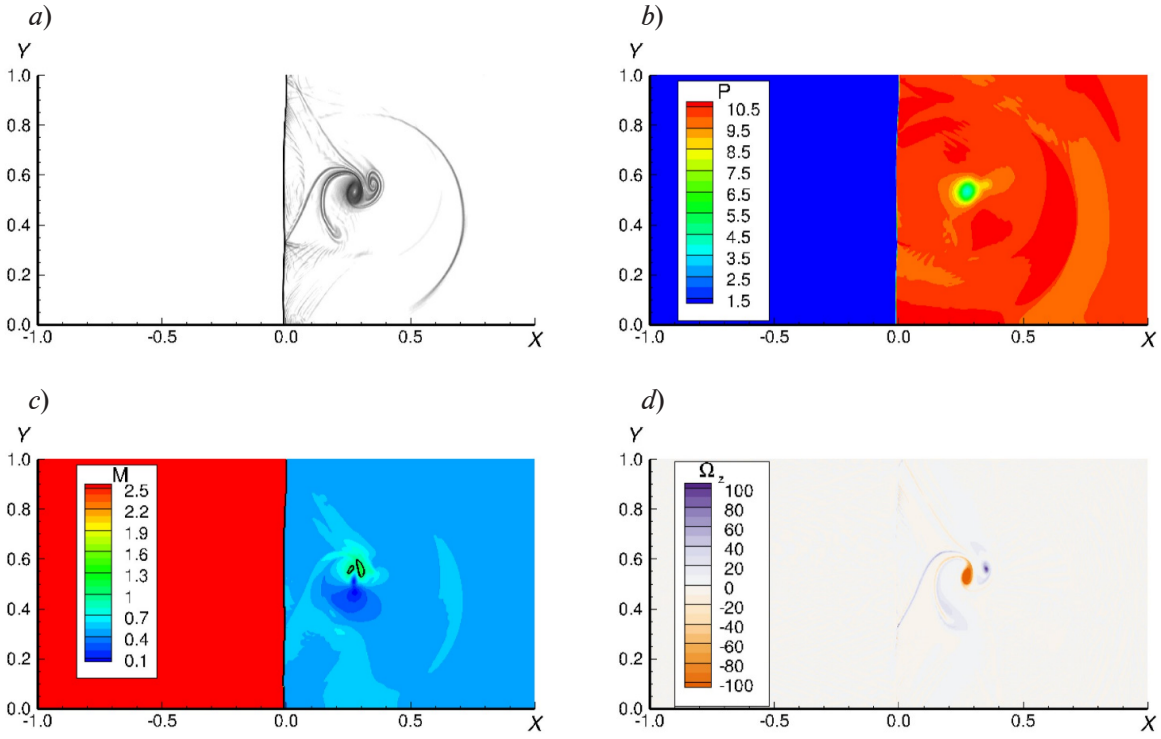

Fig. 3. Problem solution for time $t = 446.25$ ms for 1/800 grid spacing: schlieren imaging (*a*); fields of pressure, Mach number and vorticity vector projected on the z axis, respectively (*b*, *c*, *d*)

To compare the results quantitatively, we determine the integral computational error in the region Ω including a distorted vortex ($0.24 < x < 0.40$, $0.46 < y < 0.62$). The integral computational error ε is determined by the formula

$$\varepsilon = \sqrt{\frac{\sum_{i=i_1}^{i_2}\sum_{j=j_1}^{j_2}\left(\rho_{i,j} - \rho_{i,j}^{ref}\right)^2}{(i_2 - i_1)(j_2 - j_1)}} \cdot \frac{100\%}{\rho_2},$$

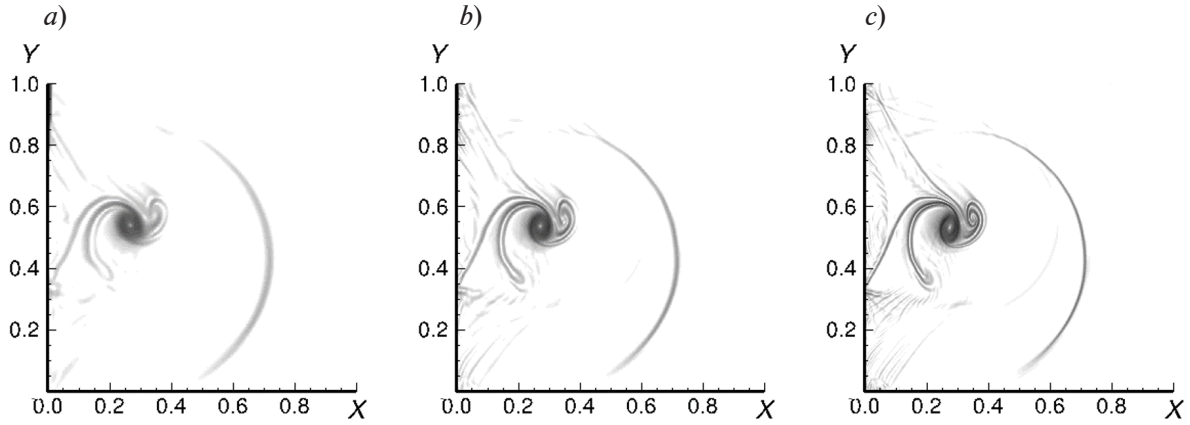

Fig. 4. Schlieren images for time t = 446.25 ms for calculation by AUSM scheme for grids of different resolutions: h =1/200 (*a*), 1/400 (*b*) and 1/800 (*c*)

where $\rho_{i,j}$, $\rho_{i,j}^{ref}$ are the average densities for the cell (i,j), respectively, in the given calculation and in the reference solution (the solution obtained by the AUSM scheme for the grid with h = 1/800 was taken as reference).

It can be seen from Table 1 (calculations by the AUSM scheme) that the integral error decreases rather quickly for more refined grids.

The second important criterion for comparing the results obtained with different grids is the integral entropy E over the region Ω, or rather its positive (E^+) and negative (E^-) components:

$$E^+ = \int_\Omega \left(\frac{\omega + |\omega|}{2} \right)^2 dxdy,\ E^- = \int_\Omega \left(\frac{\omega - |\omega|}{2} \right)^2 dxdy.$$

Table 1

Effect of grid spacing on integral characteristics in calculations using two variable reconstruction schemes

h	ε, %	E^+	E^-	δE^+	δE^-
				%	
TVD scheme					
1/200	6.2	2.5	22.7	47	31
1/400	2.3	3.9	29.4	16	11
1/500	1.3	4.2	30.9	11	6
1/600	0.3	4.4	31.8	5	4
1/800	–	4.7	33.0	–	–
WENO-5 scheme					
1/200	2.92	3.3	27.2	30	18
1/400	1.22	4.3	31.9	8	3
1/800 (TVD)	–	4.7	33.0	–	–

Notations: h is the grid spacing; ε is the integral computational error; E^+ and E^- are the positive and negative components of entropy; δE^+ and δE^- are the corresponding errors of the quantities.

A positive value corresponds to a weaker contact discontinuity, and a negative value corresponds to a stronger one. The entropies obtained in the calculations with on grids of different dimensions, the entropy values are compared with the values obtained in the reference calculation, and relative errors are calculated:

$$\delta E^+ = \frac{\left| E_h^+ - E_{h_{ref}}^+ \right|}{E_{h_{ref}}^+} \cdot 100\%,$$

$$\delta E^- = \frac{\left| E_h^- - E_{h_{ref}}^- \right|}{E_{h_{ref}}^-} \cdot 100\%.$$

Table 1 compares the integral error as well as the positive and negative components of entropy and their relative errors for solutions obtained using grids of different resolutions. Analyzing the data given in Table 1 for the calculations by the AUSM scheme, we can conclude that the absolute components of entropy grow monotonically with a decrease in grid spacing h. Furthermore, a faster drop in the relative entropy error is observed compared with the drop in integral error.

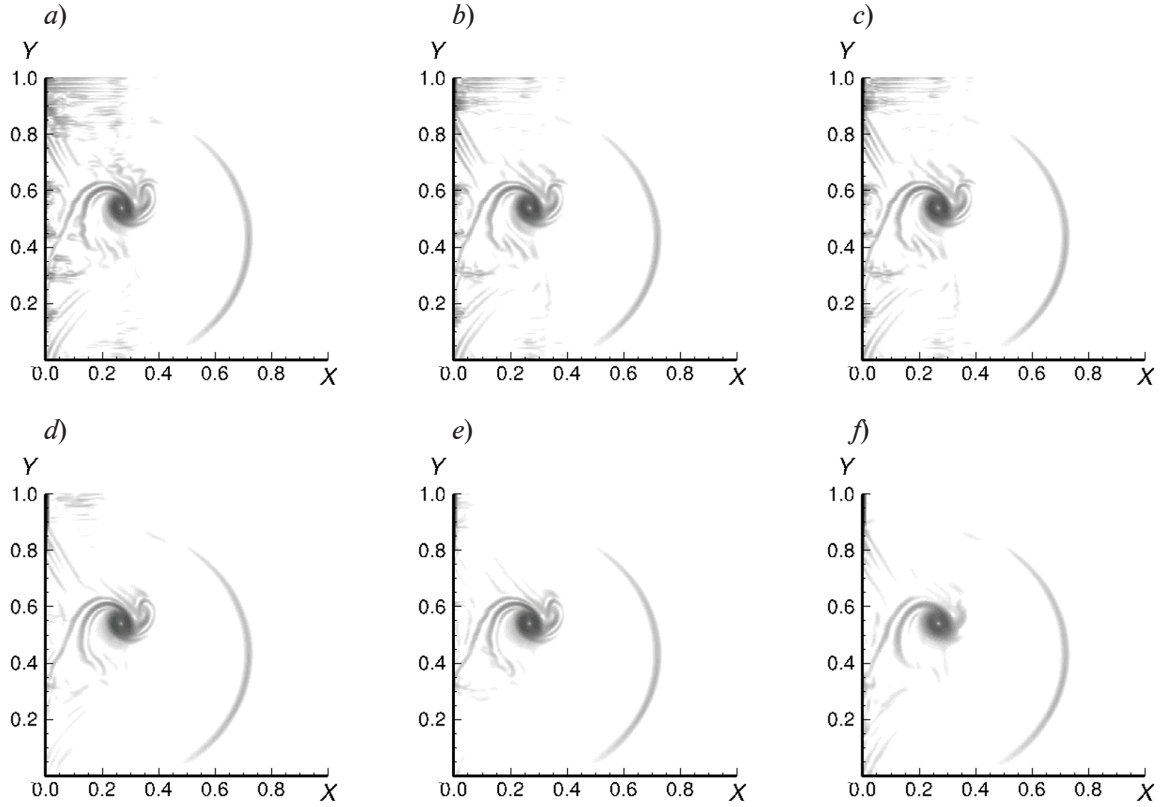

Fig. 5. Problem solution obtained on grid with spacing $h = 1/200$ using different methods for approximation of convective fluxes: Roe (*a*), HLLC (*b*), AUSMD (*c*), Godunov (*d*), AUSM (*e*), HLL (*f*)
The reconstruction of variables on the face was carried out for all cases by the TVD scheme [16] with the van Abada limiter [14]

Comparison of methods for calculating convective fluxes. Fig. 5 shows the fields of the relative density gradient for the grid spacing $h = 1/200$, constructed from the calculations using various methods for approximating convective fluxes. The reconstruction of variables at the face was carried out for all cases by the TVD scheme [16] with a van Albada limiter [14].

Comparing the integral computational error for various methods of approximating convective fluxes (Table 2, data for the TVD scheme), we find that the Roe scheme apparently yields the minimum error. However, it is worth noting that the difference in the values of the integral error ε does not exceed 1% for all considered methods. The dependence of the entropy components on the method selected turns out to be stronger: the difference between the most and least accurate schemes is 20% in terms of positive entropy and 10% in terms of negative entropy. Analysis of the field picture of the relative density gradient (see Fig. 5) indicates that more accurate solutions, obtained by schemes with lower dissipation (Roe schemes, HLLC), exhibit large numerical oscillations behind the shock wave front, whereas more dissipative schemes (HLL, AUSM) give smooth solutions. The HLL scheme turned out to be the least accurate, strongly smearing the second contact discontinuity (Fig. 5,*f*).

Comparison of schemes for reconstruction of variables with increased accuracy. High-order approximation schemes are known to have only first-order accuracy at the shock wave front [17]. However, a more accurate solution can be reconstructed in smoother sections using WENO-family schemes, which should lead, in particular, to smaller vortex dissipation in this scheme. To process the discontinuities correctly, monotonicity is preserved in this scheme family by weighing several reconstructed values taking into account smoothness indicators [14].

Table 2

Influence of method for approximation of convective fluxes on integral characteristics of problem solution

Method	ε, %	E^+	E^-	δE^+	δE^-
				%	
TVD scheme					
Roe	5.60	2.71	22.5	42	32
HLLC	5.70	2.47	22.8	47	31
AUSMD	6.06	2.51	22.6	46	32
Godunov	5.79	2.49	23.0	47	30
AUSM	6.22	2.47	22.7	47	31
HLL	6.53	1.80	19.0	62	42
AUSM (h=1/800)	–	4.69	33.0	–	–
WENO-5 scheme					
Roe	2.92	3.30	22.5	29	16
HLLC	3.02	3.40	22.8	28	16
AUSMD	2.79	3.30	22.6	29	17
Godunov	2.69	3.40	23.0	28	16
AUSM	2.98	3.30	22.7	30	18
HLL	3.56	3.00	19.0	37	20
AUSM (h=1/800)	–	4.69	33.0	–	–

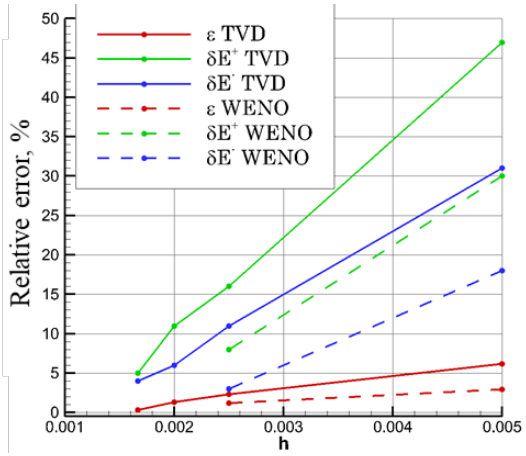

Fig. 6. Dependences of relative error on grid spacing obtained from calculations of integral characteristics, compared for different reconstruction schemes

Table 1 also shows the results obtained in calculations with grids spacings h = 1/200 and 1/400, using the WENO5 scheme for variable reconstruction. The AUSM scheme was used as an approximation scheme for convective fluxes. Comparing the integral errors obtained in the calculations using TVD and WENO schemes (see Fig. 6 and Table 2), we can see that the relative computational errors for all integral characteristics are on average half as small for the solutions obtained by the WENO scheme compared to those obtained by the TVD scheme even in the case of sufficiently coarse grids.

Fig. 7 and Table 2 give the solutions obtained with the grid spacing h = 1/200 and WENO-5 reconstruction for various schemes for approximating convective fluxes. Analyzing these results, we can conclude that significant numerical oscillations are observed behind the shock wave front in the solutions obtained by the WENO scheme, despite the better resolution. However, as in the case of TVD schemes, the solution becomes smoother when more dissipative methods are adopted for calculating convective fluxes. Therefore, combining WENO-5 and HLL (or AUSM) schemes may be a suitable option to obtain a sufficiently smooth solution to the problem, likely to be more accurate than those produced by any of the second-order TVD schemes.

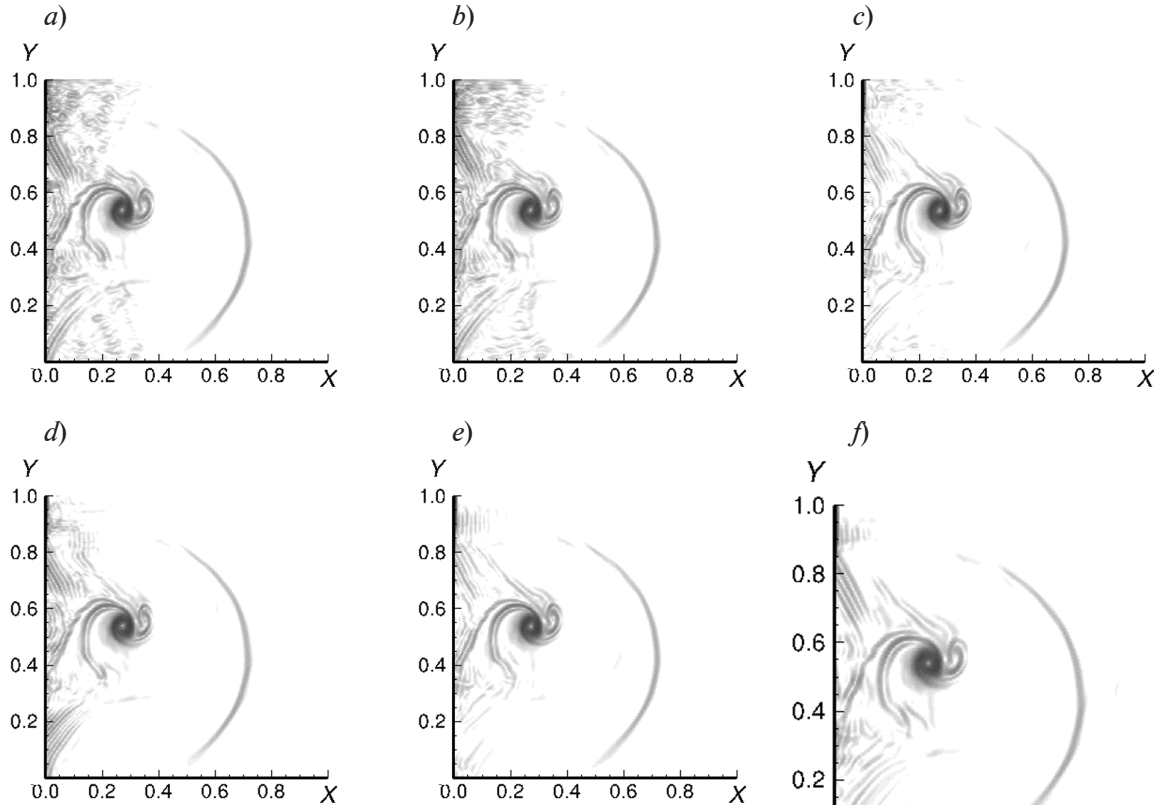

Fig. 7. Problem solution obtained on grid with spacing $h = 1/200$ using different methods for approximation of convective fluxes: Roe (*a*), HLLC (*b*), AUSMD (*c*), Godunov (*d*), AUSM (*e*), HLL (*f*)
The WENO-5 scheme was used for all cases

Conclusion

We solved a two-dimensional problem of the interaction of a shock wave and an isentropic vortex. The unstructured-grid finite-volume SINF/Flag-S code developed at Peter the Great Polytechnic University was used for the calculations. The results obtained using various schemes for approximating convective fluxes (Godunov, Roe, HLL, HLLC, AUSM, AUSMD) were compared. The influence of the scheme for variable reconstruction at the faces (second-order TVD and fifth-order WENO) was also analyzed. The problem was tested for a sequence of refined grids, where the solution for the most detailed one was considered as reference.

The weak influence of the method chosen for approximating convective fluxes on the integral computational accuracy was confirmed qualitatively and quantitatively. However, the degree of numerical oscillations in the solution varies: schemes with lower dissipation (Roe, HLLC) produce strong numerical oscillations behind the shock wave front, whereas more dissipative schemes (HLL, AUSM) yield a smooth solution.

We established a significant dependence of the computational accuracy on the variable reconstruction scheme: the solutions obtained using the WENO-5 scheme are several times more accurate than those obtained using the TVD scheme. A high level of dissipation in second-order schemes apparently leads to an artificial decrease in vorticity and less accurate resolution of the vortex structure. One of the approaches to suppressing strong numerical oscillations occurring behind the shock wave front consists of combining a high-order WENO scheme with a more dissipative method for approximating convective fluxes. With this combination, the integral accuracy of the solution is preserved but the oscillations become weaker.

REFERENCES

1. **Delery J. M.,** Aspects of vortex breakdown, Prog. Aerosp. Sci. 30 (1) (1994) 1–59.

2. **Rodionov A. V.,** Vzaimodeystviye udarnoy volny s vikhrem: testovaya zadacha dlya metodov skvoznogo scheta [The interaction between a shock wave and a vortex: a test problem for methods of through calculations]; Vychislitelnyy eksperiment v aeroakustike i aerodinamike. Sbornik tezisov devyatoy rossiyskoy konferentsii [Proceedings of the 9-th Russian Conference "Computational Experiment in Aeroacoustics and Aerodynamics"], Svetlogorsk, Kaliningrad Region, Sept. 26 – Oct. 1, 2022, Keldysh Institute of Applied Mathematics, RAS. (2022) 318–319 (in Russian).

3. **Hollingsworth M. A., Richards E. J.,** A schlieren study of the interaction between a vortex and a shock in ave in a shock tube, In book: Aeronautical Research Council. Fluid Motion Subcommittee 17985, Techn. Rep. 2323; Cranfield, UK, 1955.

4. **Dosanjh D. S., Weeks T. M.,** Interaction of a starting vortex as well as a vortex street with a traveling shock wave, AIAA J. 3 (2) (1965) 216–223.

5. **Naumann A., Hermanns E.,** On the interaction between a shock wave and a vortex field, AGARD Conf. Proc. 131 (1973) 1–23.

6. **Guichard L., Vervisch L., Domingo P.,** Two-dimensional weak shock-vortex interaction in a mixing zone, AIAA J. 33 (10) (1995) 1797–1802.

7. **Delery J., Horowitz E., Leuchter O., Solignac J.,** Fundamental studies on vortex flows, Rech. Aérospat. 2 (1984) 1–24.

8. **Kolesnik E. V., Smirnov E. M., Smirnovskiy A. A.,** Numerical solution of a 3D problem on a supersonic viscous gas flow past a plate-cylindrical body junction at M = 2.95, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 12 (2) (2019) 7–22 (in Russian).

9. **Godunov S. K.,** A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics, Mat. Sb. (N. S.) 47 (89) (3) (1959) 271–306 (in Russian).

10. **Roe P.,** Approximate Riemann solvers, parameter vectors, and difference schemes, J. Comput. Phys. 2 (43) (1981) 357–372.

11. **Harten A., Lax P. D., van Leer B.,** On upstream differencing and Godunov-type schemes for hyperbolic conservation laws, SIAM Rev. 25 (1) (1983) 35–61

12. **Toro E. F., Spruce M., Speares W.,** Restoration of the contact surface in the HLL-Riemann solver, Shock Waves. 4 (1) (1994) 25–34.

13. **Liou M.-S., Steffen C. J.,** A new flux splitting scheme, J. Comput. Phys. 107 (1) (1993) 23–39.

14. **Van Albada G. D., van Leer B., Roberts W. W.,** A comparative study of computational methods in cosmic gas dynamics, Astron. Astrophys. 108 (1) (1982) 76–84.

15. **Bakhvalov P. A., Kozubskaya T. K.,** EBR-WENO scheme for solving gas dynamics problems with discontinuities on unstructured meshes, Preprints of Keldysh Institute of Applied Mathematics, RAS, No. 23. Moscow, 2017 (in Russian).

16. **Harten A.,** High resolution schemes for hyperbolic conservation laws, J. Comput. Phys. 49 (3) (1983) 357–393.

17. **Ivanov M. Ya., Krayko A. N.,** Ob approksimatsii razryvnykh resheniy pri ispolzovanii raznostnykh skhem skvoznogo scheta [On approximation of discontinuous solutions when using difference schemes of a through calculation], Zhurnal vychislitelnoy matematiki i matematicheskoy fiziki [Journal of Numerical Mathematics and Mathematical Physics]. 18 (3) (1978) 780–783 (in Russian).

СПИСОК ЛИТЕРАТУРЫ

1. **Delery J. M.** Aspects of vortex breakdown // Progress in Aerospace Sciences. 1994. Vol. 30. No. 1. Pp. 1–59.

2. **Родионов А. В.** Взаимодействие ударной волны с вихрем: тестовая задача для методов сквозного счета // «Вычислительный эксперимент в аэроакустике и аэродинамике»: Сборник тезисов Девятой российской конференции, 26 сентября – 1 октября 2022., г. Светлогорск Калининградской области. М.: ИПМ им М. В. Келдыша. РАН, 2022. С. 318–319.

3. **Hollingsworth M. A., Richards E. J.** A schlieren study of the interaction between a vortex and a shock wave in a shock tube // Aeronautical Research Council. Fluid Motion Subcommittee 17985. Technical Report FM 2323; Cranfield, UK, 1955.

4. **Dosanjh D. S., Weeks T. M.** Interaction of a starting vortex as well as a vortex street with a traveling shock wave // AIAA (American Institute of Aeronautics and Astronautics) Journal. 1965. Vol. 3. No. 2. Pp. 216–223.

5. **Naumann A., Hermanns E.** On the interaction between a shock wave and a vortex field // AGARD Conference Proceedings. 1973. Vol. 131. Pp. 1–23.

6. **Guichard L., Vervisch L., Domingo P.** Two-dimensional weak shock-vortex interaction in a mixing zone // AIAA Journal. 1995. Vol. 33. No. 10. Pp. 1797–1802.

7. **Delery J., Horowitz E., Leuchter O., Solignac J.** Fundamental studies on vortex flows // La Recherche Aérospatiale. 1984. Vol. 2. Pp. 1–24.

8. **Колесник Е. В., Смирнов Е. М., Смирновский А. А.** Численное решение задачи обтекания цилиндрического тела, установленного на пластине, сверхзвуковым потоком вязкого газа при M = 2,95 // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2019. Т. 12. № 2. С. 7–22.

9. **Годунов С. К.** Разностный метод численного расчета разрывных решений уравнений гидродинамики // Математический сборник. 1959. Т. 47 (49). № 3. С. 271–306.

10. **Roe P.** Approximate Riemann solvers, parameter vectors, and difference schemes // / Journal of Computational Physics. 1981. Vol. 43. No. 2. Pp. 357–372.

11. **Harten A., Lax P. D., van Leer B.** On upstream differencing and Godunov-type schemes for hyperbolic conservation laws // SIAM Review. 1983. Vol. 25. No. 1. Pp. 35–61.

12. **Toro E. F., Spruce M., Speares W.** Restoration of the contact surface in the HLL-Riemann solver // Shock Waves. 1994. Vol. 4. No. 1. Pp. 25–34.

13. **Liou M.-S., Steffen C. J.** A new flux splitting scheme // Journal of Computational Physics. 1993. Vol. 107. No. 1. Pp. 23–39.

14. **Van Albada G. D., van Leer B., Roberts W. W.** A comparative study of computational methods in cosmic gas dynamics // Astronomy & Astrophysics. 1982. Vol. 108. No. 1. Pp. 76–84.

15. **Бахвалов П. А., Козубская Т. К.** Схема EBR-WENO для решения задач газовой динамики с разрывами на неструктурированных сетках. Препринты ИПМ им. М. В. Келдыша. № 23. Москва, 2017. 32 с.

16. **Harten A.** High resolution schemes for hyperbolic conservation laws // Journal of Computational Physics. 1983. Vol. 49. No. 3. Pp. 357–393.

17. **Иванов М. Я., Крайко А. Н.** Об аппроксимации разрывных решений при использовании разностных схем сквозного счета // Журнал вычислительной математики и математической физики. 1978. Т. 18. No. 3. С. 780–783.

THE AUTHORS

BABICH Elena V.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
lll.helen.lll@mail.ru
ORCID: 0000-0003-1442-4886

KOLESNIK Elizaveta V.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
kolesnik_ev@mail.ru
ORCID: 0000-0002-6308-733X

СВЕДЕНИЯ ОБ АВТОРАХ

БАБИЧ Елена Викторовна — *студентка Физико-механического института Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
lll.helen.lll@mail.ru
ORCID: 0000-0003-1442-4886

КОЛЕСНИК Елизавета Владимировна — *кандидат физико-математических наук, доцент Высшей школы прикладной математики и вычислительной физики Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
kolesnik_ev@mail.ru
ORCID: 0000-0002-6308-733X

Received 06.02.2024. Approved after reviewing 16.02.2024. Accepted 16.02.2024.
Статья поступила в редакцию 06.02.2024. Одобрена после рецензирования 16.02.2024. Принята 16.02.2024.

Original article
DOI: https://doi.org/10.18721/JPM.17202

AN ALGORITHM FOR FINDING THE DAMPING COEFFICIENTS BASED ON VIBRATION SURVEYS USING THE FREQUENCY DOMAIN DECOMPOSITION (FDD) METHOD

E. A. Andrianova✉

"B. E. Vedeneev VNIIG", JSC, St. Petersburg, Russia

✉ elena-andrianova@bk.ru

Abstract. The article presents an algorithm and a theoretical justification for the method of finding the damping coefficients on the base of vibration surveys using the frequency domain decomposition (FDD) method. This method is used for monitoring the structures such as buildings, bridges, dams, to determine experimentally their state (under operating conditions) without application of vibroexciters.

Keywords: frequency domain decomposition, damping coefficient, spectral density matrix, natural mode shape and frequency

Citation: Andrianova E. A., An algorithm for finding the damping coefficients based on vibration surveys using the frequency domain decomposition method (FDD) method, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 21–32. DOI: https://doi.org/10.18721/JPM.17202

Научная статья
УДК
DOI: https://doi.org/10.18721/JPM.17202

АЛГОРИТМ НАХОЖДЕНИЯ КОЭФФИЦИЕНТОВ ДЕМПФИРОВАНИЯ ПО ДАННЫМ ВИБРАЦИОННЫХ ОБСЛЕДОВАНИЙ МЕТОДОМ FDD (ДЕКОМПОЗИЦИИ В ЧАСТОТНОЙ ОБЛАСТИ)

Е. А. Андрианова✉

АО «ВНИИГ им. Б. Е. Веденеева», Санкт-Петербург, Россия

✉ elena-andrianova@bk.ru

Аннотация. В статье представлен алгоритм и теоретическое обоснование методики нахождения коэффициентов демпфирования по данным вибрационных обследований при использовании метода декомпозиции в частотной области (FDD). Этот метод применяется при динамическом тестировании сооружений (здания, мосты, плотины) для экспериментального определения их динамических характеристик в условиях нормальной эксплуатации без применения вибровозбудительного оборудования.

Ключевые слова: декомпозиция в частотной области, коэффициент демпфирования, спектральная плотность сигнала, собственные частоты, формы собственных колебаний

Ссылка для цитирования: Андрианова Е. А. Алгоритм нахождения коэффициентов демпфирования по данным вибрационных обследований методом FDD (декомпозиции в частотной области) // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 2. С. 21–32. DOI: https://doi.org/10.18721/JPM.17202

Introduction

Full-scale dynamic testing of building structures remains a crucial problem, especially for unique structures, such as dams. Experimental assessment of dynamic characteristics (natural frequencies, mode shapes, relative damping ratios) and monitoring of these characteristics allows to control the safety, strength, integrity of the structure as well as to identify substantial changes without resorting to specialized devices or visual inspection of each structure.

Operational modal analysis (OMA) comprises an entire group of methods aimed at experimentally determining the dynamic characteristics of structures under normal operating conditions. These methods have become increasingly widespread for diagnostics of dynamic characteristics in various structures as advances are made in measuring and recording systems. A particular popular method of the OMA group is Frequency Domain Decomposition (FDD) [2–4]. The FDD method and the ARTeMIS Modal software based on it have been adopted since 2019 by the scientists of the B.E. Vedeneev All-Russian Research Institute of Hydraulic Engineering (St. Petersburg, Russia) [9].

In addition to determining natural frequencies and mode shapes, the EFDD method expanding the capabilities of FDD also offers an algorithm for determining damping ratios [4–6], however, it is rather complex, often yielding large errors.

A simpler, more accurate algorithm is proposed in this paper for identifying the damping parameters.

Our goal was to formulate and theoretically substantiate a new method for finding damping ratios based on vibration surveys.

The FDD method is described in detail in [2, 3, 7], and its theoretical framework is formulated in [1]. The algorithm of this method consists of the following mandatory steps.

Step 1. A cross-spectral density matrix (CSDM) $\mathbf{G}_y(\omega)$ of simultaneously measured vibration signals is calculated for each frequency ω of a given range.

Step 2. A singular-value decomposition (SVD) of the matrices $\mathbf{G}_y(\omega)$ is performed at each frequency ω, their first singular value $\sigma_1(\omega)$ is determined, and a frequency function of the first singular value is constructed, averaged over all measurements.

Note that the main idea of the FDD algorithm (see, for example, [2, 3, 7]) is that the first singular value $\sigma_1(\omega)$ of the matrix $\mathbf{G}_y(\omega_m)$ has local maxima near modal frequencies. The mathematical justification for this was given in [1]. Alternatively, we use the function $\sigma_1(\omega)$ to determine the logarithmic decrements corresponding to each natural frequency.

Theoretical justification of the procedure for determining logarithmic decrements

The response $\mathbf{y}(t)$ of the system is uniquely decomposed into their linear combination (due to the linear independence of the eigenmodes):

$$\mathbf{y}(t) = \boldsymbol{\varphi}_1 \cdot q_1(t) + \boldsymbol{\varphi}_2 \cdot q_2(t) + \ldots = \boldsymbol{\Phi}\, \mathbf{q}(t). \tag{1}$$

As found in [7], if white noise is considered as external force, and dissipation is assumed to be small, the following expression holds true for CSDM $\mathbf{G}_y(\omega)$:

$$\mathbf{G}_y(\omega) = \sum_{m=1}^{M} \frac{c_m\, \boldsymbol{\varphi}_m \boldsymbol{\varphi}_m^{\mathrm{H}}}{i\omega - \lambda_m} + \frac{c_m\, \boldsymbol{\varphi}_m \boldsymbol{\varphi}_m^{\mathrm{H}}}{-i\omega - \lambda_m^*} = \boldsymbol{\Phi} \cdot \mathrm{diag}\left(2\,\mathrm{Re}\left(\frac{c_m}{i\omega - \lambda_m} \right) \right) \cdot \boldsymbol{\Phi}^{\mathrm{H}}, \tag{2}$$

where λ_m is the pole,

$$\lambda_m = -\gamma_m + i\omega_{dm}, \tag{3}$$

$$(\gamma_m = \omega_{0m} \varsigma_m); \tag{4}$$

$\boldsymbol{\varphi}_m$ is the eigenmode; $\boldsymbol{\Phi}$ is a matrix whose columns are eigenvectors, $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots \boldsymbol{\varphi}_M]$; M is the number of modes accounted for in decomposition (1); c_m is a positive coefficient; i is the imaginary unit; H is the Hermitian conjugate.

The component ω_{dm} of expression (3) is the natural frequency accounting for damping. In expression (4), ω_{0m} is the natural frequency without accounting for damping; ς_m is the damping ratio.

Next, we introduce the notation

$$\alpha_m(\omega) = 2\,\mathrm{Re}\left(\frac{c_m}{i\omega - \lambda_m}\right) = \frac{c_m \gamma_m}{(\omega - \omega_{md})^2 + \gamma^2{}_m}. \tag{5}$$

Notably, modal vectors $\boldsymbol{\varphi}_m$ in expression (2) are assumed to be normalized, since the coefficient α_m contains, according to expression (5), a constant c_m that can be supplemented with a normalization factor.

Then expression (2) can be written as follows:

$$\mathbf{G}_y(\omega) = \boldsymbol{\Phi} \cdot \mathrm{diag}\left(\alpha_m(\omega)\right) \cdot \boldsymbol{\Phi}^{\mathrm{H}}, \tag{6}$$

or

$$\mathbf{G}_y(\omega) = \sum_{m=1}^{M} \alpha_m \boldsymbol{\varphi}_m \boldsymbol{\varphi}_m^{\mathrm{H}}. \tag{7}$$

Dynamic testing of the structure can be used to calculate the CSDM of the measured signals over a certain frequency range; next, the SVD of these matrices can be used to obtain the frequency function of the first singular value [10–12].

It was proved in our earlier study [1] that the natural frequencies of the structure considered are located near the local maxima of this function. This paper proposes and substantiates a method for determining the damping ratios based on an experimentally constructed function of the first singular value.

An obvious idea is to compare the analytical expression for the first singular value and the experimentally constructed function [14]. Even though there is no general analytical expression

for the first singular value of CSDM, a fairly good analytical approximation can be obtained under certain conditions.

Consider two main cases when it is possible to obtain such an approximation.

Case of single natural frequency. Here, the response $\mathbf{y}(t)$ of the system (see Eq. (1)) in the vicinity of some natural frequency (let us denote it as ω_{ds}) is determined mainly by eigenmode with the same number. Then the following relation holds true:

$$\mathbf{y}(t) \approx \boldsymbol{\varphi}_s \cdot q_s(t), \tag{8}$$

and the expression for the CSDM given by Eq. (4) can be rewritten as:

$$\mathbf{G}_y(\omega) \approx \alpha_s \boldsymbol{\varphi}_s \boldsymbol{\varphi}_s^{\mathrm{H}}. \tag{9}$$

Evidently, Eq. (9) holds true when the values of the functions $\alpha_s(\omega)$ significantly exceed the other values $\alpha_m(\omega)$ in the vicinity of the frequency ω_{ds}. Now let us find the conditions under which this requirement is satisfied.

Consider the properties of functions $\alpha_m(\omega)$. It can be seen from expression (5) that these functions depend on the corresponding natural frequencies and damping ratios. They have one maximum each, reached at the corresponding natural frequency.

Indeed (see our study [1]), determining the extreme values of the functions $\alpha_m(\omega)$, we obtain for $\omega = \omega_{dm}$

$$\alpha_m(\omega_{dm}) = \mathrm{c}_m / \gamma_m. \tag{10}$$

We introduce the notation for the minimum distance d_s on the frequency scale between the frequency ω_{ds} and the rest of the natural frequencies ω_{dm}, namely:

$$d_s = \min_{m \neq s} \left| \omega_{ds} - \omega_{dm} \right|. \tag{11}$$

Then the following relation holds true for the frequency ω_{dm} with all $m \neq s$:

$$\alpha_m(\omega_{ds}) \leq \frac{c_m \gamma_m}{d_s^2 + \gamma_m^2} = \frac{c_m / \gamma_m}{(d_s / \gamma_m)^2 + 1}. \tag{12}$$

Now let us introduce the notation

$$r_{sm} = d_s / \gamma_m = \frac{d_s}{\omega_{0m} \cdot \varsigma_m}. \tag{13}$$

Comparing relations (12) and (10), we can see that the condition

$$\alpha_s(\omega_{ds}) >> \alpha_m(\omega_{ds}) \tag{14}$$

is satisfied if

$$r_{sm} >> 1, \text{ for all } s \neq m. \tag{15}$$

Therefore, as condition (15) is satisfied, only one term in expression (7) can be taken into account, i.e., the CSDM is calculated by Eq. (9).

An analytical expression can be constructed for the first singular value pf the CSDM $\mathbf{G}_y(\omega)$ described by expression (9). The matrix $\mathbf{G}_y(\omega)$ is evidently square and symmetrical (it is Hermitian for complex modes). The rank of the matrix $\mathbf{G}_y(\omega)$ equals unity (since the rank of the product of the matrices does not exceed the ranks of the multipliers), therefore this matrix has no more than one eigenvalue other than zero.

We find it by defining the eigenvalues. Let u be the eigenvector and λ the eigenvalue of the matrix $\mathbf{G}_y(\omega)$; then, by defining the eigenvector and the eigenvalue, we obtain the following equality:

$$\mathbf{G}_y(\omega) u = \lambda u, \tag{16}$$

then,

$$\alpha_s \boldsymbol{\varphi}_s \boldsymbol{\varphi}_s^{\mathrm{H}} u = \alpha_s \boldsymbol{\varphi}_s (\boldsymbol{\varphi}_s^{\mathrm{H}} u) = \alpha_s (\boldsymbol{\varphi}_s^{\mathrm{H}} u) \boldsymbol{\varphi}_s = \lambda u. \tag{17}$$

An immediate consequence of equality (17) is that the only nonzero eigenvector $u = \boldsymbol{\varphi}_s$, and the eigenvalue $\lambda = \alpha_s \left\| \boldsymbol{\varphi}_s \right\|^2$. Evidently, $\lambda \geq 0$, since the coefficient $c_m \geq 0$. Consequently, the matrix $\mathbf{G}_y(\omega)$ is positive semi-definite, and then (since it is also Hermitian), its singular values coincide with its eigenvalues. Since $\left\| \boldsymbol{\varphi}_s \right\|^2 = 1$, then, apparently, σ_1 coincides with α_s. Therefore, the maximum singular value of the CSDM in the vicinity of natural frequencies can be written as

$$\sigma_1 = \alpha_s = \frac{c_s \gamma_s}{(\omega - \omega_{ds})^2 + \gamma_s^2} = \frac{c_s \gamma_s}{(\omega - \omega_{ds})^2 + \omega_{0s}^2 \cdot \varsigma_s^2}. \tag{18}$$

Furthermore, the natural frequencies accounting for damping (ω_{dk}) and without it (ω_{0k}) practically coincide for small damping ratios.

If we compare the function of the first singular value, obtained by processing experimental data, with analytical dependence (18), we can estimate the logarithmic decrements. Let us rewrite Eq. (18) in the following form:

$$\sigma_1 = \frac{A}{(\omega - \omega_{ds})^2 + \omega_{ds}^2 B^2}. \tag{19}$$

Then we use the least squares method (for example), making it possible to determine the coefficients A and B that approximate the analytical function σ_1 as close as possible (see Eq. (19)) to the experimentally obtained dependence in the vicinity of some natural frequency. The value of the parameter B apparently corresponds to the damping decrement.

The case of two converged natural frequencies. Let us now consider the second case, when the values of two natural frequencies numbered k and $k + 1$ (ω_{dk} and ω_{dk+1}) are located close to each other (such frequencies are generally known as converged in the literature), i.e., condition (15) is not satisfied for frequencies with these numbers.

However, if condition (15) is satisfied for all other natural frequencies, except for frequencies numbered k and $k + 1$, then CDSM can be calculated by Eq. (7) with only two terms:

$$\mathbf{G}_y(\omega) = \sum_{m=k}^{k+1} \alpha_m \boldsymbol{\varphi}_m \boldsymbol{\varphi}_m^{\mathrm{H}}, \tag{20}$$

or the matrix has the coordinate form

$$\mathbf{G}_y = \begin{bmatrix} \sum\limits_{m=k}^{k+1} \alpha_m (\varphi_m^{(1)})^2 & \sum\limits_{m=k}^{k+1} \alpha_m \varphi_m^{(1)} \varphi_m^{(2)} & & \sum\limits_{m=k}^{k+1} \alpha_m \varphi_m^{(1)} \varphi_m^{(N)} \\ \sum\limits_{m=k}^{k+1} \alpha_m \varphi_m^{(2)} \varphi_m^{(1)} & \sum\limits_{m=k}^{k+1} \alpha_m (\varphi_m^{(2)})^2 & \cdots & \sum\limits_{m=k}^{k+1} \alpha_m \varphi_m^{(2)} \varphi_m^{(N)} \\ \cdots & \cdots & & \cdots \\ \sum\limits_{m=k}^{k+1} \alpha_m \varphi_m^{(N)} \varphi_m^{(1)} & \sum\limits_{m=k}^{k+1} \alpha_m \varphi_m^{(N)} \varphi_m^{(2)} & & \sum\limits_{m=k}^{k+1} \alpha_m (\varphi_m^{(N)})^2 \end{bmatrix}. \tag{21}$$

This matrix has the dimensions $N \times N$, and it is difficult to find its singular values (or eigenvalues) analytically.

To simplify the task, we compose a Gram matrix (denoting it as $\mathbf{K}$) based on the vectors $\sqrt{\alpha_k}\boldsymbol{\varphi}_k$ and $\sqrt{\alpha_{k+1}}\boldsymbol{\varphi}_{k+1}$.

The matrix $\mathbf{K}$ has the following coordinate form:

$$\mathbf{K} = \begin{bmatrix} \alpha_k & \sqrt{\alpha_k \alpha_{k+1}}(\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_{k+1}) \\ \sqrt{\alpha_k \alpha_{k+1}}(\boldsymbol{\varphi}_{k+1}, \boldsymbol{\varphi}_k) & \alpha_{k+1} \end{bmatrix}. \tag{22}$$

This matrix, like the matrix $\mathbf{G}_y$, is Hermitian and positive semi-definite (a property of the Gram matrix). A proof that nonzero eigenvalues of the matrix $\mathbf{G}_y$ coincide with the eigenvalues of the matrix $\mathbf{K}$ (the Gram matrix constructed from the corresponding vectors) was given in [1], and the eigenvalues of the matrices $\mathbf{K}$ and $\mathbf{G}_y$ coincide with their singular values. Thus, the first singular value of the matrix $\mathbf{G}_y$ is equal to the spectral radius of the matrix $\mathbf{K}$.

In this case, the matrix $\mathbf{K}$ has the dimensions of 2×2, and we can easily construct an analytical expression for the value of its spectral radius [16]:

$$\rho = \frac{\mathrm{tr}(\mathbf{K}) + \sqrt{\mathrm{tr}^2(\mathbf{K}) - 4\det(\mathbf{K})}}{2}. \tag{23}$$

Since it follows from the expression for matrix (22) that

$$\det(\mathbf{K}) = \alpha_k \alpha_{k+1} - \alpha_k \alpha_{k+1} \left|(\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_{k+1})\right|^2, \tag{24}$$

$$\mathrm{tr}(\mathbf{K}) = \alpha_k + \alpha_{k+1}, \tag{25}$$

we obtain the following expression for the spectral radius of the matrix **K** (and therefore, for the first singular value of the matrix $\mathbf{G}_y$):

$$\sigma_1 = \frac{\alpha_k + \alpha_{k+1} + \sqrt{(\alpha_k - \alpha_{k+1})^2 + 4\alpha_k \alpha_{k+1} \left|(\varphi_k, \varphi_{k+1})\right|^2}}{2}. \tag{26}$$

It follows directly from Eq. (26) that if φ_k is orthogonal to φ_{k+1}, then $\sigma_1 = \max\limits_{i=k,k+1} \alpha_i$, regardless of whether condition (15) is satisfied.

Fig. 1 shows graphs of functions σ_1 and α_i calculated for a system with three degrees of freedom. Evidently, in the vicinity of the first natural frequency ω_1, function of the first singular value σ_1 coincides well with the function α_1 in a fairly wide frequency range ($r_{12} = 12.7$ in this example). However, the behavior of the function σ_1 changes for the second and third natural frequencies ($r_{23} = 0.85, r_{32} = 1.7$, respectively), while the greatest difference between the graph of this function and the corresponding graphs of functions α_i is observed in the frequency range corresponding to the interval between the maxima of the curves for the functions α_2 and α_3.

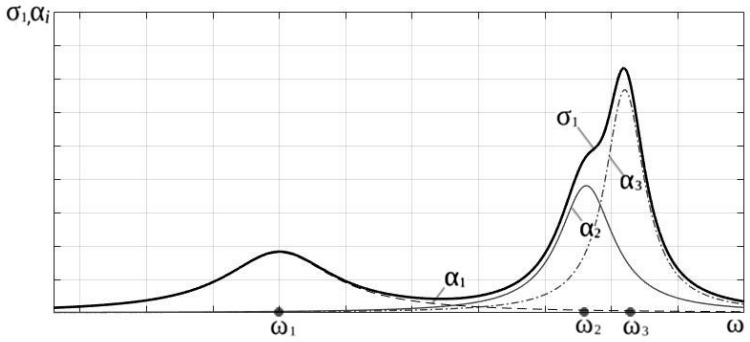

Fig. 1. Behavior of first singular value σ_1 compared with functions $\alpha_i(\omega)$ for system with three degrees of freedom ($\omega_1 - \omega_3$)

Note that expression (26) is simplified at the point of intersection of the curves for α_k and α_{k+1} (α_2 and α_3 in the example):

$$\sigma_1 = \alpha_k (1 + \left|(\varphi_k, \varphi_{k+1})\right|). \tag{27}$$

In other cases, expression (26) as a function of logarithmic decrements is a rather complex expression, so using it to find the necessary parameters turns out to be a difficult task.

We propose a different approach to solving this problem in the case of converged frequencies. It is known from linear algebra that the sum of the eigenvalues of a square matrix is equal to its trace [15–17], and since, as noted above, the eigenvalues and singular values of the matrix **K** coincide in this case, the following formula holds true:

$$\sigma_1 + \sigma_2 = \alpha_k + \alpha_{k+1}. \tag{28}$$

Let us introduce the notation

$$s(\omega) = \sigma_1(\omega) + \sigma_2(\omega). \tag{29}$$

Then expression (28) can be rewritten as follows:

$$s = \alpha_k + \alpha_{k+1} \tag{30}$$

(for brevity, the argument is omitted here).

If we substitute the expressions for α_k and α_{k+1} in Eq. (30), then the following relation holds true:

$$s = \frac{c_k \gamma_k}{(\omega - \omega_{dk})^2 + \omega_{0k} \varsigma_k^{\,2}} + \frac{c_{k+1} \gamma_{k+1}}{(\omega - \omega_{dk+1})^2 + \omega_{0k+1} \varsigma_{k+1}^{\,2}}. \quad (31)$$

Fig. 2 shows a comparison of the sum of the first two singular values s and the sum $\alpha_k + \alpha_{k+1}$ for the previously considered case (see Fig. 1), a system with three degrees of freedom ($\alpha_2 + \alpha_3$ in the example).

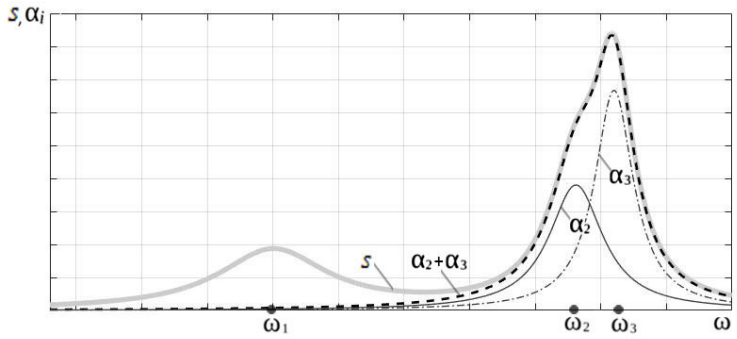

Fig. 2. Sum of first two singular values $s(\omega)$ compared with functions $\alpha_i(\omega)$ and sum $\alpha_k + \alpha_{k+1}$ for system with three degrees of freedom (see Fig. 1)

Singular-value decomposition of the matrix $\mathbf{G}_y(\omega)$ is performed during processing of experimental data obtained by dynamic measurements at each frequency ω of a given range; not only the first singular value $\sigma_1(\omega)$ but also the remaining singular values of the function are determined. Thus, the sum of the first two singular values is known.

Similar to the case of a single natural frequency, the natural frequencies accounting for damping (ω_{dk}) and without it (ω_{0k}) practically coincide for small damping ratios. Therefore, the analytical expression for the sum s of the first two singular values has the form

$$s = \frac{A}{(\omega - \omega_{dk})^2 + \omega_{dk}^{\,2} B^2} + \frac{C}{(\omega - \omega_{dk+1})^2 + \omega_{dk+1}^{\,2} D^2}, \quad (32)$$

where *A, B, C, D* are unknown parameters.

The least squares method can be used to determine these unknown parameters. Evidently, the coefficient *B* is an estimate of the damping ratio ς_k, and the coefficient *D* corresponds to the ratio ς_{k+1}.

Example calculations of damping ratios

The method for determining logarithmic decrements was tested with a mathematical model of a system with 8 degrees of freedom (Fig. 3).

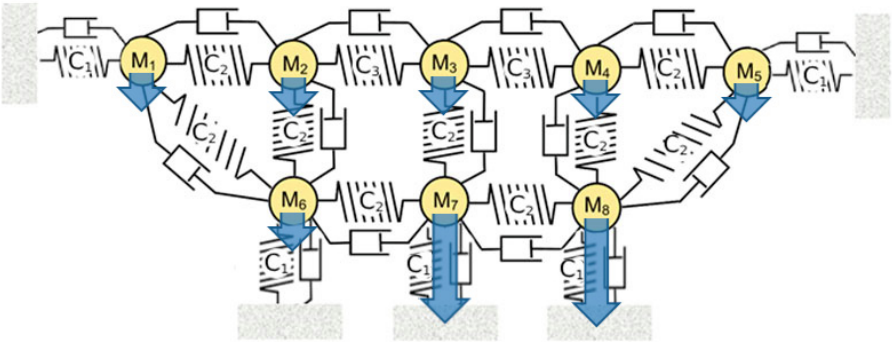

Fig. 3. Model of dynamic system with 8 degrees of freedom

Table 1

Parameters set for model of dynamic system with 8 degrees of freedom and their values (see Fig. 3)

Mass, kg	Stiffness, N/m	Force amplitude, N (applied to mass)
$M_1 = \ldots = M_8 = 25.9$	$C_1 = 770$, $C_2 = 1000$, $C_3 = 950$	$F_1 = \ldots = F_6 = 1$, $F_7 = F_8 = 10$

Note. The damping ratios (logarithmic decrements) were set to be the same and equal to 0.01.

The inertial and stiffness parameters were set in the adopted model. The damping ratios were assumed to be the same and equal to 0.01. Then, a proportional damping matrix was calculated at each natural frequency by the specified damping parameters. Forces with a white noise spectrum were applied to masses $M_1 - M_8$. The values of the force amplitude and other parameters of the system are given in Table 1. The loading modeled was non-uniformly distributed over the degrees of freedom: the amplitude of the force at degrees of freedom 7 and 8 was increased by 10 times.

Next, vibrational responses at all degrees of freedom were determined as time series with a given frequency from the exact solution of the dynamic problem.

The measurement data obtained by this approach were used to test the FDD technique and to subsequently identify the damping parameters by the proposed method. The results were compared with the parameters set for the model. The frequency function of the first singular value is shown in Fig. 4. Evidently, the six peaks corresponding to the natural frequencies (1–5 and 8) can be regarded as, i.e., the problem of determining the damping ratios corresponds to the case of a single natural frequency. The damping ratios corresponding to frequencies 1–5 and 8 were calculated by Eq. (19) based on to the algorithm described above (Table 2, upper lines).

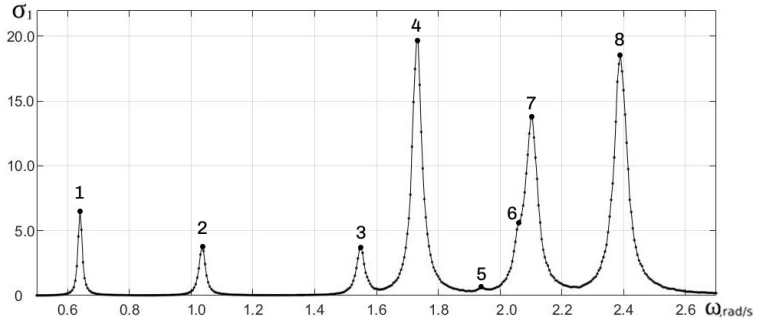

Fig. 4. Frequency spectrum for function of first singular value in CSDM for system with 8 degrees of freedom (see Fig. 3)

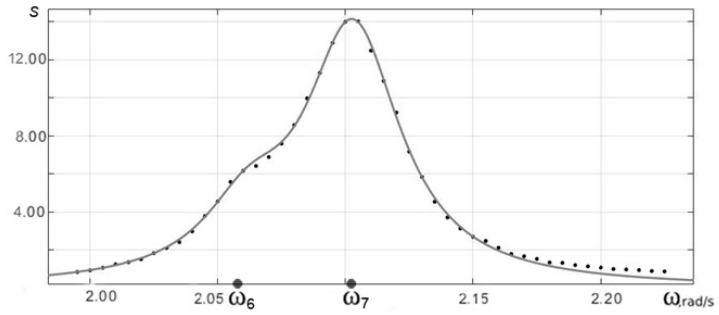

Fig. 5. Calculated sum of first two singular values (points) and obtained dependence approximated by Eq. (32) (solid line

Table 2

Calculated damping ratios and corresponding coefficients of determination

Peak	Natural frequency	Damping ratio	Coefficient of determination
Case of 'single' natural frequency			
1	0.64	0.0107	0.9985
2	1.03	0.0109	0.9995
3	1.54	0.0104	0.9957
4	1.73	0.0104	0.9975
5	1.93	0.0140	0.9357
8	2.39	0.0106	0.9983
Case of two 'converged' natural frequencies			
6	2.06	0.01014	0.9977
7	2.10	0.01070	0.9981

Note. The least squares method was used to find the values of the coefficient of determination.

The graph in Fig. 5 for the case of converged frequencies (these are frequencies 6 and 7 in the example) shows the points corresponding to the sum of the first and second singular values (obtained by the simulation model) as well as a function approximating them by Eq. (32).

The values of the identified damping ratios and the corresponding coefficients of determination found by the above method are also given in Table 2.)

Conclusion

The paper proposes a simple method for determining the damping ratios after identifying the natural frequencies of a structure based on experimental data using the FDD technique. Analytical expressions are obtained in the vicinity of natural frequencies for the first singular value as well as for the sum of the first two singular values as frequency functions. The method for determining the damping ratios is based on approximating the values obtained from processing the experimental data by analytical expressions with unknown parameters. The least squares method allows to determine the damping ratios. The first singular value is approximated in the case of a single natural frequency, the sum of the first two singular values is approximated in the case of converged natural frequencies.

The damping ratios were identified for a model problem. The proposed method has an advantage over the well-known EFDD method presented in [4], due to its lower complexity; furthermore, unlike the EFDD method, it allows determining the damping characteristics in the case of natural frequencies that are close in value.

REFERENCES

1. **Andrianova E. A., Novitskii D. V., Onischuk V. S.,** Theoretical justification of natural frequency identification in the FDD (Frequency Domain Decomposition) method, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 16 (1) (2023) 82–96 (in Russian).

2. **Brincker R., Zhang L., Andersen P.,** Modal identification from ambient responses using frequency domain decomposition, In: Proc. 18th Int. Modal Analysis Conf. (IMAC 18), San Antonio, Texas, USA, Febr. 7–10 (2000) 625–630.

3. **Brincker R., Zhang L., Andersen P.,** Modal identification of output-only systems using frequency domain decomposition, Smart Mater. Struct. 10 (3) (2001) 441–445.

4. **Brincker R., Ventura C. E., Andersen P.,** Damping estimation by frequency domain decomposition, In: Proc. 19th Int. Modal Analysis Conf. (IMAC 19), Kissimmee, Florida, USA, Febr. 5–8 (2001) 698–703.

5. **Zhang L., Tamura Y.,** Damping estimation of engineering structures with ambient response measurements, In: Proc. 21st Int. Modal Analysis Conf. & Expos. 2003 (IMAC XXI) "A Conf. & Expos. on Structural Dynamics". In 3 Vols. Vol. 1. Kissimmee, Florida, USA, Febr. 3–6, 2003. (2003) 1023–1029.

6. **Brincker R., Zhang L.,** Frequency domain decomposition revisited, In: Proc. 3rd Int. Operational Modal Analysis Conf. (IOMAC), Portonovo (Ancona), Italy, May 4–6, 2009. (2009) 615–626.

7. **Brincker R., Ventura C.,** Introduction to operational modal analysis, John Wiley and Sons, Ltd., Hoboken, New Jersey, USA, 2015.

8. **Rainieri C., Fabbrocino G.,** Operational modal analysis of civil engineering structures: An introduction and guide for applications, Springer, New York, 2014.

9. **Amador S., Juul M., Friis T., Brincker R.,** Application of Frequency Domain Decomposition Identification technique to half spectral densities, In book: "Topics in Modal Analysis & Testing". Vol. 9. Conf. Proc. of the Soc. for Experimental Mechanics. Ed. by M. Mains, B. J. Dilworth. Ch. 39. The Society for Experimental Mechanics, Inc. Springer, Cham, New York (2019) 343–346.

10. ARTeMIS Modal. Software for Operational Modal Analysis and Experimental Modal Analysis. URL: http://www.svibs.com/. Date of last access: 10.12.2023.

11. **Hasan M. D. A., Ahmad Z. A. B., Leong M. S., Hee L. M.,** Enhanced frequency domain decomposition algorithm: A review of a recent development for unbiased damping ratio estimates, J. Vibroengineering. 20 (5) (2018) 1919–1936.

12. **Lang G. F.,** Matrix madness and complex confusion. Review of complex modes from multiple viewpoints, Sound Vib. (November – Structural Analysis) (2012) 8–12.

13. **Tarpo P., Olsen P., Amador S., et al.,** On minimizing the influence of the noise tail of correlation functions in operational modal analysis, Proc. Eng. 199 (2017) 1038–1043.

14. **Kravchenko K. I., Kugaevsky S. S., Zhuravlyov M. P., Elkind D. M.,** Detection of natural oscillation frequencies and coefficients of relative damping of dynamic system at milling by method of operation modal analysis, Bulletin of the Bryansk State Technical University. (8 (61)) (2017) 14–24 (in Russian).

15. **Ilyin V. A., Poznyak E. G.,** Linear algebra, Victor Kamkin, 1987.

16. **Karchevskiy E. M., Karchevskiy M. M.,** Lektsii po lineynoy algebre i analiticheskoy geometrii [Lectures on linear algebra and analytical geometry], Lan' Publishing, St. Petersburg, 2018 (in Russian).

17. **Fomin V. I.,** On the lower linear bound of the spectral radius for a normal matrix, Tambov University Reports. Series: Natural and Technical Sciences. 6 (2) (2001) 145–146 (in Russian).

СПИСОК ЛИТЕРАТУРЫ

1. **Андрианова Е. А., Новицкий Д. В., Онищук В. С.** Теоретическое обоснование идентификации собственных частот в методе FDD (декомпозиции в частотной области) // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 16. № 1. С. 82–96.

2. **Brincker R., Zhang L., Andersen P.** Modal identification from ambient responses using frequency domain decomposition // Proceedings of the 18th International Modal Analysis Conference (IMAC 18). San Antonio, Texas, USA, February 7–10, 2000. Pp. 625–630.

3. **Brincker R., Zhang L., Andersen P.** Modal identification of output-only systems using frequency domain decomposition // Smart Materials and Structures. 2001. Vol. 10. No. 3. Pp. 441–445.

4. **Brincker R., Ventura C. E., Andersen P.** Damping estimation by frequency domain decomposition // Proceedings of the 19th International Modal Analysis Conference (IMAC 19), Kissimmee, Florida, USA. February 5–8, 2001. Pp. 698–703.

5. **Zhang L., Tamura Y.** Damping estimation of engineering structures with ambient response measurements // Proceedings of the 21st International Modal Analysis Conference & Exposition (IMAC XXI) "A Conference & Exposition on Structural Dynamics". In 3 Vols. Vol. 1. Kissimmee, Florida, USA, February 3–6, 2003. Pp. 1023–1029.

6. **Brincker R., Zhang L.** Frequency domain decomposition revisited // Proceedings of the 3rd International Operational Modal Analysis Conference (IOMAC), Portonovo (Ancona), Italy. May 4–6, 2009. Vol. 1. Pp. 615–626.

7. **Brincker R., Ventura C.** Introduction to Operational Modal Analysis. Hoboken, New Jersey, USA: John Wiley and Sons, Ltd., 2015. 360 p.

8. **Rainieri C., Fabbrocino G.** Operational Modal Analysis of civil engineering structures: An introduction and guide for applications. New York: Springer, 2014. 322 p.

9. **Amador S., Juul M., Friis T., Brincker R.** Application of Frequency Domain Decomposition Identification technique to half spectral densities // Topics in Modal Analysis & Testing. Vol. 9. Conference Proceedings of the Society for Experimental Mechanics. Edited by M. Mains, B. J. Dilworth. Chapter 39. The Society for Experimental Mechanics, Inc. New York: Springer, Cham, 2019. Pp. 343–346.

10. ARTeMIS Modal. Software for operational modal analysis and experimental modal analysis. Режим доступа: http://www.svibs.com/. Дата последнего обращения: 10.12.2023.

11. **Hasan M. D. A., Ahmad Z. A. B., Leong M. S., Hee L. M.** Enhanced frequency domain decomposition algorithm: A review of a recent development for unbiased damping ratio estimates // Journal of Vibroengineering. 2018. Vol. 20. No. 5. Pp. 1919–1936.

12. **Lang G. F.** Matrix madness and complex confusion. Review of complex modes from multiple viewpoints // Sound & Vibrations. 2012. November – Structural Analysis. Pp. 8–12.

13. **Tarpo P., Olsen P., Amador S., Juul M., Brincker R.** On minimizing the influence of the noise tail of correlation functions in operational modal analysis // Procedia Engineering. 2017. Vol. 199. Pp. 1038–1043.

14. **Кравченко К. Ю., Кугаевский С. С., Журавлев М. П., Элькинд Д. М.** Выявление собственных частот колебаний и коэффициентов относительного демпфирования динамической системы при фрезеровании методом операционного модального анализа // Вестник Брянского государственного технического университета. 2017. № 8 (61). С. 14–24.

15. **Ильин В. А., Позняк Э. Г.** Линейная алгебра. 6-е изд., стер. М.: Физматлит, 2007. 280 с.

16. **Карчевский Е. М., Карчевский М. М.** Лекции по линейной алгебре и аналитической геометрии. 2-е изд. СПб.: Изд-во «Лань», 2018. 424 с.

17. **Фомин В. И.** О нижних линейных оценках спектрального радиуса нормальной матрицы // Вестник Тамбовского университета. Серия: Естественные и технические науки. 2001. Т. 6. № 2. С. 145–146.

THE AUTHOR

ANDRIANOVA Elena A.
"B. E. Vedeneev VNIIG", JSC
21 Gzhatskaya St., Saint Petersburg, 195220, Russia
elena-andrianova@bk.ru
ORCID: 0000-0001-9199-6875

СВЕДЕНИЯ ОБ АВТОРЕ

АНДРИАНОВА Елена Александровна – *научный сотрудник акционерного общества «Всероссийский научно-исследовательский институт гидротехники (ВНИИГ) им. Б. Е. Веденеева».*
195220, Россия, г. Санкт-Петербург, Гжатская ул., 21
elena-andrianova@bk.ru
ORCID: 0000-0001-9199-6875

Received 07.12.2023. Approved after reviewing 07.02.2024. Accepted 07.02.2024.
Статья поступила в редакцию 07.12.2023. Одобрена после рецензирования 07.02.2024. Принята 07.02.2024.

Original article
DOI: https://doi.org/10.18721/JPM.17203

THE WAYS TO FORM FULLERENS STRUCTURE EXEMPLIFIED BY C_{28} AND C_{30} ISOMERS

A. N. Matvienko✉

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ matvienko_an@spbstu.ru

Abstract. The systematization of possible ways to form C_{28} and C_{30} fullerene isomers has been completed in this work. An additional confirmation was obtained that fullerenes created by incorporating dimers into the initial structure had a minimum energy. Two other methods for the formation of new fullerenes were also considered, namely, the connection of two domes with the same symmetry and that of fullerenes with compatible symmetry. The symmetry order of the studied fullerenes varied from the second to the seventh ones. A collaboration analysis of the formation energy values and geometric simulation results allowed us to draw conclusions about the structural changes in the resulting fullerenes.

Keywords: dimer, isomer, symmetry, fullerene, energy

Citation: Matvienko A. N., The ways to form fullerens structure exemplified by C_{28} and C_{30} isomers, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 33–38. DOI: https://doi.org/10.18721/JPM.17203

Научная статья
УДК 54.022
DOI: https://doi.org/10.18721/JPM.17203

СПОСОБЫ ФОРМИРОВАНИЯ СТРУКТУРЫ ФУЛЛЕРЕНОВ НА ПРИМЕРЕ ИЗОМЕРОВ C_{28} И C_{30}

А. Н. Матвиенко✉

Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉ matvienko_an@spbstu.ru

Аннотация. В работе завершена систематизация возможных путей образования изомеров фуллеренов C_{28} и C_{30}. Получено дополнительное подтверждение, что фуллерены, образованные встраиванием димеров в исходную структуру, имеют минимальную энергию. Рассмотрены и два других способа образования новых фуллеренов, а именно – соединение двух куполов с одинаковой симметрией и соединение фуллеренов с совместимой симметрией. Порядок симметрии исследуемых фуллеренов меняется от второго к седьмому. Совместный анализ значений энергии формирования и результатов геометрического моделирования фуллеренов позволил сделать выводы о структурных изменениях полученных структур.

Ключевые слова: димер, изомер, симметрия, фуллерен, энергия

Ссылка для цитирования: Матвиенко А. Н. Способы формирования структуры фуллеренов на примере изомеров C_{28} и C_{30} // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 2 № .17. С. 33–38. DOI: https://doi.org/10.18721/JPM.17203

Introduction

A classification of the most natural mechanisms for the formation of new fullerenes was carried out in [1]. These are fusion of fullerenes with compatible symmetry, embedding of a carbon dimer into the initial fullerene and fusion of carbon domes with the same symmetry [2–5].

This study concludes the consideration of structures and energies of C_{28} and C_{30} fullerenes formed by these mechanisms.

The goal of the study is to examine a series of fullerene isomers from C_{20} to C_{30}, using them as reference for verifying the particular mechanisms (discovered earlier) for formation of new fullerenes.

Structures of fullerenes with single (simple) and double bonds were analyzed for this purpose. The criterion for the presence of double bonds was symmetrical arrangement of their maximum number.

Formation of C_{28} fullerene isomers

Let us consider five main methods for obtaining new fullerene models.

Embedding of a dimer into a C_{26} polyhedron

Perfect fullerenes

Method I. Embedding a dimer into the hexagon from the front, parallel to the third-order axis, turns fullerene C_{26} into fullerene C_{28}. A perfect fullerene with second-order rotation/reflection symmetry is formed. Such a fullerene contains twelve pentagons and four hexagons.

Method II. A dimer is embedded at a 60° angle to the third-order symmetry axis of the initial fullerene. A perfect fullerene with second-order rotation/reflection symmetry is obtained. The initial fullerene should be rotated by 90° to obtain a mirror image in case I, and only by 30° in case II. The resulting fullerene also contains twelve pentagons and four hexagons (Fig. 1).

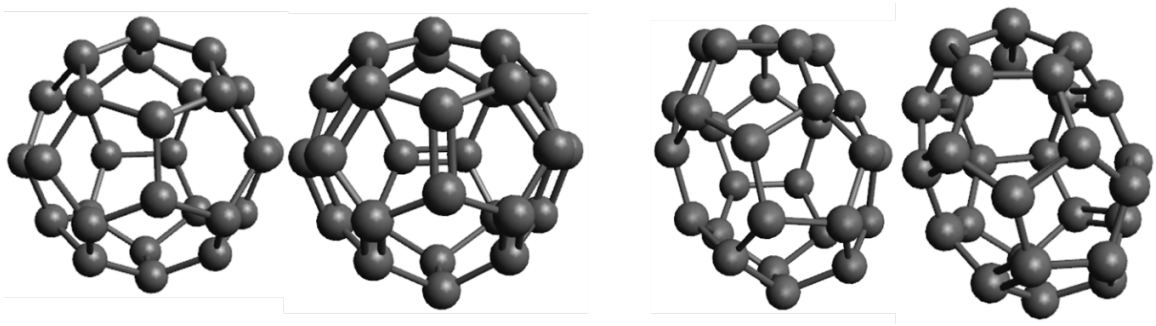

Fig. 1. Embedding of carbon dimer into initial perfect fullerene; methods with parallel embedding (method I) and embedding at a 60° angle (method II) are shown

Method III. Embedding of a dimer into a C_{26} polyhedron. Topological symmetry. Embedding a dimer into a hexagon in the background transforms the fullerene C_{24} into the fullerene C_{26}. Since it contains two additional interstices, an imperfect fullerene with topological symmetry of the third order is obtained [6]. The resulting fullerene contains two triangles, six pentagons and eight hexagons.

Method IV. Fusion of two different domes with compatible symmetry. One of the possible techniques is fusion of the C_{10} dome with the C_{18} bowl. Such a reaction can be written in the following form:

$$C_{10} + C_{18} \rightarrow C_{28}.$$

Method V. **Fusion of two heptagonal prisms.** Both initial fullerenes have the same seventh-order symmetry. The reaction has the following form (Fig. 2):

$$C_{14} + C_{14} \rightarrow (C_{14}C_{14}) \rightarrow C_{28}.$$

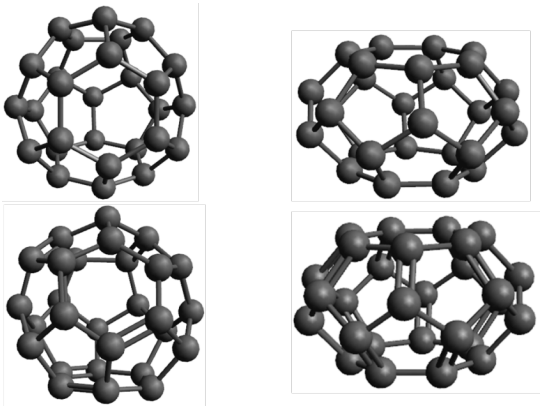

Fig. 2. Fusion of two different domes (method IV) and two heptagonal prisms (method V)

Formation of C_{30} fullerene isomers

Two angular methods for embedding the dimer. Perfect fullerenes

Method VI. A carbon dimer is embedded into perfect fullerene C_{28} [7] (see method I) in one of two hexagons at the equator of the initial fullerene. Fullerene C_{28} is transformed to fullerene C_{30}. The polyhedron formed contains twelve pentagons and five hexagons.

Method VII. The initial configuration is shown in Fig. 1 (see method II). The initial half of the fullerene must be rotated by 90° to obtain a mirror image in the previous case (see method VI). In this case, rotation by only 30° is sufficient. There is another difference: the embedding is performed at a 60° angle. Such operations transform the initial fullerene into fullerene C_{30}. The resulting fullerene also contains twelve pentagons and five hexagons (Fig. 3) [8].

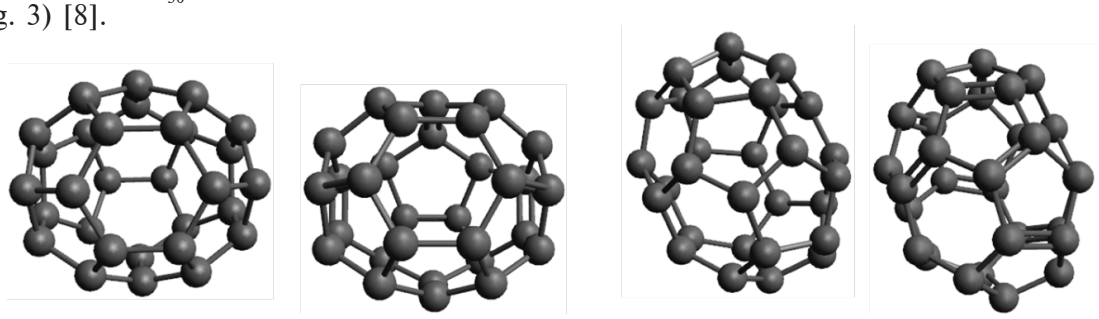

Fig. 3. Two angular methods for dimer embedding: at the equator (method VI) and at 60° angle (method VII)

Method VIII. Fusion of dome C_{10} with dome C_{20}. Perfect fullerene. Both initial configurations have fifth-order symmetry. A fullerene consisting of five squares, two pentagons and ten hexagons is formed after fusion. A perfect fullerene with fifth-order symmetry is produced.

Method IX. Fusion of two C_{15} domes. Perfect fullerene. Both domes have fifth-order symmetry. The resulting fullerene has twelve pentagons and five hexagons. The obtained structure is a perfect fullerene with fifth-order symmetry (Fig. 4) [9].

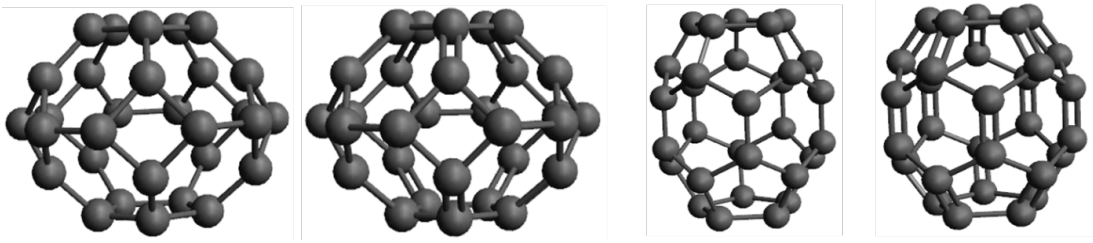

Fig. 4. Fusion of different (method VIII) and identical (method IX) domes

Obtained isomers and their formation energy

Energy calculations for the considered isomers were performed in the Avogadro editor. The table shows the calculated values of the minimum energy for fullerenes with single bonds (E_{min}) and the maximum energy for fullerenes with double bonds (E_{max}).

Table

Calculated formation energies of fullerene isomers

Chemical formula	Method	Energy, kJ/mol	
		E_{min}	E_{max}
C_{28}	I	924	1790
	II	**689**	1872
	III	2250	3867
	IV	924	1172
	V	1119	2264
C_{30}	VI	**832**	847
	VII	**689**	1872
	VIII	1840	2548
	IX	**852**	1907

Practically all isomers presented in Table are (with the exception of those formed by method III) perfect fullerenes. The analysis of the obtained results shows that fullerenes formed by embedding a dimer into the initial fullerenes [5] have minimal energies (highlighted in bold). The energy of fullerene formation increases when different or identical domes are fused. The fullerenes considered have second, third, fifth and seventh order symmetry.

Conclusion

The study concludes the systematization of possible methods for forming chains of fullerene isomers from C_{20} to C_{30}, specifically, C_{28} and C_{30}. This systematization allowed to obtain additional evidence confirming the conclusions drawn in the previous work [10]. It was found earlier that formation of fullerenes by embedding a dimer in the initial fullerene is energetically advantageous. The resulting fullerenes are perfect. An exception is the case when the initial fullerene contains two additional interstices. Such a fullerene turns out to be imperfect with topological symmetry, taking high energy costs to produce.

Our findings provide a deeper understanding of fullerene structure. From a practical standpoint, these data can be useful to researchers choosing a specific fullerene with a focus on its formation energy in a wide variety of projects.

REFERENCES

1. **Melker A. I., Matvienko A. N.,** Periodic system of fullerenes: isomers from C_{20} to C_{28}, In: Proc. NDTCS-2019: Nano-Design, Technology, Computer Simulations (24.09.2019 – 27.09.2019. Brest, Belarus) 72–78.

2. **Poklonski N. A., Ratkevich S. V., Vyrko S. A.,** Quantum-chemical calculation of carbodo-decahedron formation in carbon plasma, J. Phys. Chem. A. 119 (34) (2015) 9133–9139.

3. **Sánchez-Bernabe F. J.,** Towards a periodic pattern in classical and nonclassical fullerenes with tetrahedral structure, Mater. Physic. Mech. 45 (1) (2020) 79–86.

4. **Melker A. I., Krupina M. A., Matvienko A. N.,** Nucleation and growth of fullerenes and nanotubes having four-fold symmetry, Mater. Phys. Mech. 47 (2) (2021) 315–343.

5. **Endo M., Kroto H. W.,** Formation of carbon nanofibers, J. Phys. Chem. 96 (17) (1992) 6941–6943.

6. **Schwerdtfeger P., Wirz L. N., Avery J.,** The topology of fullerenes, WIREs Comput. Mol. Sci. 5 (1) (2015) 96–145.

7. **Tunuguntla R. H., Henley R. Y., Yao Y.-C., et al.,** Enhanced water permeability and tunable ion selectivity in subnanometer carbon nanotube porins, Sci. 357 (6353) (2017) 792–796.

8. **Amaya T., Natata T., Hirao T.,** Synthesis of highly strained π-bowls from sumanene, J. Am. Chem. Soc. 131 (31) (2009) 10810–10811.

9. **Amiri H., Shepard K. L., Nickolls C., Sánchez R. H.,** Single-walled carbon nanotubes: Mimics of biological channels, Nano Lett. 17 (2) (2017) 1204–1211.

10. **Matvienko A. N.,** The ways to form midi-fullerens structure exemplified by C_{22}, C_{24} and C_{26} isomers, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 16 (1) (2023) 125–131 (in Russian).

СПИСОК ЛИТЕРАТУРЫ

1. **Melker A. I., Matvienko A. N.,** Periodic system of fullerenes: isomers from C_{20} to C_{28} // Proceedings of NDTCS-2019: Nano-Design, Technology, Computer Simulations. 24.09.2019 – 27.09.2019. Brest, Belarus. Pp.72–78.

2. **Poklonski N. A., Ratkevich S. V., Vyrko S. A.** Quantum-chemical calculation of carbodo-decahedron formation in carbon plasma // The Journal of Physical Chemistry. A. 2015. Vol. 119. No. 34. Pp. 9133–9139.

3. **Sánchez-Bernabe F. J.** Towards a periodic pattern in classical and nonclassical fullerenes with tetrahedral structure // Materials Physics and Mechanics. 2020. Vol. 45. No. 1. Pp. 79–86.

4. **Melker A. I., Krupina M. A., Matvienko A. N.** Nucleation and growth of fullerenes and nanotubes having four-fold symmetry // Materials Physics and Mechanics. 2021. Vol. 47. No. 2. Pp. 315–343.

5. **Endo M., Kroto H. W.** Formation of carbon nanofibers // The Journal of Physical Chemistry. 1992. Vol. 96. No. 17. Pp. 6941–6943.

6. **Schwerdtfeger P., Wirz L. N., Avery J.** The topology of fullerenes // WIREs Computational Molecular Science. 2015. Vol. 5. No. 1. Pp. 96–145.

7. **Tunuguntla R. H., Henley R. Y., Yao Y.-C., Pham T. A., Wanunu M., Noy A.** Enhanced water permeability and tunable ion selectivity in subnanometer carbon nanotube porins // Science. 2017. Vol. 357. No. 6353. Pp. 792–796.

8. **Amaya T., Natata T., Hirao T.** Synthesis of highly strained π-bowls from sumanene // Journal of the American Chemical Society. 2009. Vol. 131. No. 31. Pp. 10810–10811.

9. **Amiri H., Shepard K. L., Nickolls C., Sánchez R. H.** Single-walled carbon nanotubes: Mimics of biological channels // Nano Letters. 2017. Vol. 17. No. 2. Pp. 1204–1211.

10. **Матвиенко А. Н.** Способы формирования структуры миди-фуллеренов на примере изомеров C_{22}, C_{24} и C_{26} // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 16. № 1. С. 125–131.

THE AUTHOR

MATVIENKO Aleksandra N.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
matvienko_an@spbstu.ru
ORCID: 0000-0002-3012-1407

СВЕДЕНИЯ ОБ АВТОРЕ

МАТВИЕНКО Александра Николаевна – *ассистент Высшей школы механики и процессов управления*; *инженер той же школы Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
matvienko_an@spbstu.ru
ORCID: 0000-0002-3012-1407

Received 17.04.2023. Approved after reviewing 29.01.2024. Accepted 29.01.2024.
Статья поступила в редакцию 17.04.2023. Одобрена после рецензирования 29.01.2024. Принята 29.01.2024.

MATHEMATICAL PHYSICS

Original article
DOI: https://doi.org/10.18721/JPM.17204

ON ASSOCIATED HOMOGENEOUS GELFAND FUNCTIONS

A. S. Berdnikov[1], *A. L. Bulyanitsa*[1,2]✉, *K. V. Solovyev*[2,1]

[1] Institute for Analytical Instrumentation RAS, St. Petersburg, Russia;

[2] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉antbulyan@yandex.ru

Abstract. The paper proposes refined definitions for associated homogeneous functions (AHFs) of real variables, which are of great practical importance for a wide range of problems. I. M. Gelfand and Z. Ya. Shapiro were the first in 1955 to introduce AHFs into scientific use. However, the possibilities of using these functions in various applications have not been exhausted to this day. The proposed definitions inherit the basic idea of the original paper to define chains of new functions using the recurrent linear functional relations, where some homogeneous Euler function is the starting point. This makes it possible to apply the corresponding results not only for differentiable and continuous functions, but also for discontinuous functions, including discontinuous ones at all points. The possibility of constructing a detailed consistent theory of AHFs of real variables, defined by a chain of linear recurrent functional relations of a general form, is shown. The basic theorems are formulated and proven. Further ways of generalizing the functions under consideration, are discussed.

Keywords: associated homogeneous Gelfand functions, homogeneous Euler function, recurrent linear functional relations

Funding: The reported study was carried out at Institute for Analytical Instrumentation, RAS (St. Petersburg, Russia) within the framework of the State Assignment of the Ministry of Science and Higher Education of the Russian Federation, No. 075-00439-24-00, dated 27.12.2023.

Citation: Berdnikov A. S., Bulyanitsa A. L., Solovyev K. V., On associated homogeneous Gelfand functions, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 39–60. DOI: https://doi.org/10.18721/JPM.17204

Научная статья
УДК 517.982.42
DOI: https://doi.org/10.18721/JPM.17204

О ПРИСОЕДИНЕННЫХ ОДНОРОДНЫХ ФУНКЦИЯХ ГЕЛЬФАНДА

А. С. Бердников[1], *А. Л. Буляница*[1,2]✉, *К. В. Соловьев*[2,1]

[1] Институт аналитического приборостроения РАН, Санкт-Петербург, Россия;

[2] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉ antbulyan@yandex.ru

Аннотация. В статье предлагаются уточненные определения для присоединенных однородных функций (ПОФ) вещественных переменных, имеющих большое прикладное значение для широкого круга задач. Понятие ПОФ было впервые сформулировано И. М. Гельфандом и З. Я. Шапиро в 1955 году, но возможности использования этих

функций в разнообразных приложениях не исчерпаны и поныне. Предлагаемые определения наследуют базовую идею оригинальной статьи: определять цепочки новых функций с помощью рекуррентных линейных функциональных соотношений, начиная с некоторой одиночной однородной функции Эйлера; это позволяет использовать соответствующие результаты не только для дифференцируемых и непрерывных функций, но и для разрывных, в том числе разрывных во всех точках. Показана возможность построения развернутой непротиворечивой теории ПОФ вещественных переменных, определяемых с помощью цепочки линейных рекуррентных функциональных соотношений общего вида. Формулируются и доказываются базовые теоремы теории рассматриваемых функций. Обсуждаются дальнейшие пути обобщения указанного класса функций.

Ключевые слова: присоединенные однородные функции Гельфанда, однородные функции Эйлера, рекуррентные линейные функциональные соотношения

Финансирование. Работа выполнена в Институте аналитического приборостроения Российской академии наук (Санкт-Петербург) при частичной поддержке в рамках Государственного задания Министерства науки и высшего образования Российской Федерации № 00-24-00439-075 от 27.12.2023 г.

Ссылка для цитирования: Бердников А. С., Буляница А. Л., Соловьев К. В. О присоединенных однородных функциях Гельфанда // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 2. С. 39–60. DOI: https://doi.org/10.18721/JPM.17204

Introduction

A fundamental study by Gelfand and Shapiro published in [1] introduced, along with other important categories, the concept of infinite chains of associated homogeneous functions (ASFs), also used by the authors to construct a special kind of generalized functions. This concept was later used almost unchanged in monograph [2]. This was perhaps the first significant advance made in the general theory of homogeneous functions (see §§187, 188 in [3]) since Euler's time[1].

Like many seemingly abstract mathematical constructs, homogeneous Euler functions and, consequently, associated homogeneous Gelfand functions turn out to be a useful tool for solving a variety of applied problems, including those far removed from the original goals intended by the authors. Such applications include the similarity principle of nonrelativistic trajectories of charged particles in Euler-homogeneous electrostatic fields, proposed by Golikov [5–10]. It is proved in the above studies that the trajectory of a charged particle in the given electrostatic field is scaled as a single entity with proper scaling of initial coordinates and initial kinetic energy (preserving the initial exit angles). Practical examples of this effective tool for synthesis of various electron and ion-optical systems can be found in [11–21].

The new idea proposed in [1] was quickly developed by other mathematicians (see, for example, [22–33]). Unfortunately, unlike the original study [1], where differentiation with respect to degree of homogeneity was used to construct a chain of functions, the associated homogeneous functions in subsequent publications are determined axiomatically using binomial functional recurrence relations[2]. However, with this approach, researchers actually deal only with a narrow subset of Gelfand functions instead of the supposed infinite chain of associated homogeneous functions f_k (namely, the cases $k = 0$ and $k = 1$, see below). As a result, without serious revision of definitions and proofs the statements to be proved are generally speaking only valid for this narrow subset.

[1] In this regard, it is also worth mentioning the theory of spherical functions that are homogeneous harmonic functions of three variables [4], developed in the 19th century.

[2] This approach to defining associated homogeneous functions is mentioned in [1, 2] as an alternative technique but is not used in practice, see Appendix

However, this important remark does not apply to the results obtained in the original study by Gelfand and Shapiro. In our opinion, the logic by which a new class of functions of real variables[3] is defined in [1, 2] is as follows (the corresponding reasoning is presented more accurately in the Appendix).

Homogeneous Euler functions of degree p satisfy the functional relation

$$\forall \lambda \in R, \forall \mathbf{x} \in R^n : \; f_p\left(\lambda x_1, \lambda x_2, \ldots, \lambda x_n\right) = \lambda^p f_p\left(x_1, x_2, \ldots, x_n\right). \tag{1}$$

If $x_1 > 0$, these functions can be represented in the most general form as

$$f_p\left(x_1, x_2, \ldots, x_n\right) = \left(x_1\right)^p h\left(\frac{x_2}{x_1}, \frac{x_3}{x_1}, \ldots \frac{x_n}{x_1}\right), \tag{2}$$

where h is an arbitrary function ($n = 1$) of real variables (see §§187, 188 in book [3]).

Performing k-fold differentiation of homogeneous functions (2) with respect to the parameter p (homogeneity degree of the function), we obtain functions $f_{p,k}$ with log-periodic power-law singularities. Indeed, in view of Eq. (2), if $x_1 > 0$ and $k \geq 0$, these functions can be represented with an accuracy up to an auxiliary multiplier $1/k!$ as

$$f_{p,k}\left(x_1, x_2, \ldots, x_n\right) = \frac{1}{k!}\left(x_1\right)^p \left(\ln x_1\right)^k h\left(\frac{x_2}{x_1}, \frac{x_3}{x_1}, \ldots \frac{x_n}{x_1}\right), \tag{3}$$

where k is a non-negative integer (the order of the function)[4].

Ref. [1] primarily concentrates on studying the integrals of the form

$$\int_0^b x^p \ln^k x \cdot \varphi\left(x\right) dx, \; \int_{-a}^0 \left|x\right|^p \ln^k \left|x\right| \cdot \varphi\left(x\right) dx,$$

where $\varphi(x)$ are functions without singularities, including at zero, which can be differentiated with respect to the variable x a sufficient number of times.

Furthermore, it is important to study similar multidimensional integral expressions. Evidently, if $p \leq -1$, the given integrals diverge. However, it is proved in [1] using the analytical continuation procedure and the framework of generalized function theory that these integrals can be given a meaningful numerical value, thus defining a new class of generalized functions.

Functions of form (3) comprise the basis for functions with log-periodic power-law singularities, serving to solve the problem under consideration with increased efficiency. Differentiation with respect to the parameter p as well as expansion into Laurent series and into power series with respect to the parameter p serve in [1] (and especially in monograph [2]) as the main tool for constructing new generalized functions, whose values are associated with these integrals.

Chains of functions $f_{p,k}$ are of undoubted interest in themselves. The authors of [1] call the functions $f_{p,k}$ associated homogeneous, by analogy with associated eigenvectors of linear operators[5], and differentiation with respect to the parameter p is used as the main tool.

[3] The presentation in this paper is focused on functions of real variables, although all of these results are of course also largely valid for functions of complex variables.

[4] Note that the differentiation procedure requires caution, since the function h in Eq. (2) may depend explicitly or implicitly on the homogeneity degree p. It also requires a reasoned transition from a fixed homogeneity degree to a continuously variable parameter which is differentiable. In particular, the homogeneity degree in Eq. (2) can be considered as function $\omega(p)$ of the abstract parameter p with respect to which differentiation is performed, and the identity $\omega(p) \equiv p$, generally speaking, is not guaranteed.

[5] Homogeneous Euler functions of degree p can be considered as eigenfunctions of the linear scaling operator $L[f]$: $f(\mathbf{x}) \rightarrow f(\lambda \mathbf{x})$ corresponding to the eigenvalues of λ^p.

However, along with defining associated homogeneous functions as a result of differentiation of homogeneous functions with respect to homogeneity degree p, Refs. [1, 2] also formulate the assumption (not quite accurate and not particularly significant for the authors' further reasoning) that chains of functions $f_{p,k}$ can be regarded as partial solutions of the following linear functional recurrence relations:

$$\begin{aligned} &\text{for } k=0,\ f_{p,0}(\lambda\mathbf{x})=a(\lambda)f_{p,0}(\mathbf{x}),\\ &\text{for } k\geq 1,\ f_{p,k}(\lambda\mathbf{x})=a(\lambda)f_{p,k}(\mathbf{x})+b(\lambda)f_{p,k-1}(\mathbf{x}), \end{aligned} \tag{4}$$

with the coefficients $a(\lambda)=\lambda^p,\ b(\lambda)=\lambda^p\ln\lambda$.

It follows from analysis of relations (4) given in [1, 2] for particular cases $k = 0$ and $k = 1$ that the functions

$$a(\lambda)=\lambda^p,\ b(\lambda)=c\lambda^p\ln\lambda,$$

where $c \neq 0$ is an arbitrary constant, are the only situation where nontrivial continuous solutions are possible (but not necessarily exist) for system (4).

Without loss of generality, a value $c = 1$ can be used, since the parameter $c \neq 0$ is reduced upon substitution of $f_{p,k}(\mathbf{x})\to c^k f^*_{p,k}(\mathbf{x})$. Apparently, the chains of functions $f_{p,k}$ of form (3) with the subscript k shifted by $m \geq 0$, which are padded with zeros, namely,

$$f^{(m)}_{p,k}(x_1,x_2,\ldots,x_n)=\begin{cases}\dfrac{1}{(k-m)!}(x_1)^p(\ln x_1)^{k-m}\,h_m\left(\dfrac{x_2}{x_1},\dfrac{x_3}{x_1},\ldots\dfrac{x_n}{x_1}\right) & \text{for } k\geq m,\\ 0 & \text{for } k<m,\end{cases} \tag{5}$$

while their linear combinations[6] are also solutions of functional relations (4), if such a solution is a chain of functions $f_{p,k}$.

The authors of [1, 2] subsequently confined the consideration to functions of form (3), since more complex constructs are not required to study integrals with log-periodic power-law singularities and give these integrals reasonable meaning at negative degrees.

As mentioned earlier, unlike the original studies [1, 2], subsequent works by other authors [22–33] used binomial functional relations (4), where $a(\lambda)=\lambda^p$, $b(\lambda)=\lambda^p\ln\lambda$, as a formal definition of associated homogeneous functions. Unfortunately, however, direct substitution shows, in particular, that associated homogeneous functions (3), obtained from homogeneous functions (2) by repeated differentiation with respect to the parameter p (as done in [1, 2]), are not solutions to functional relations (4) even when $k \geq 2$:

$$\begin{aligned} &f_{p,0}(\lambda\mathbf{x})=\lambda^p f_{p,0}(\mathbf{x}),\\ &f_{p,1}(\lambda\mathbf{x})=\lambda^p\ln\lambda f_{p,0}(\mathbf{x})+\lambda^p f_{p,1}(\mathbf{x}),\\ &f_{p,2}(\lambda\mathbf{x})=\frac{\lambda^p\ln^2\lambda}{2}f_{p,0}(\mathbf{x})+\lambda^p\ln\lambda f_{p,1}(\mathbf{x})+\lambda^p f_{p,2}(\mathbf{x}),\\ &f_{p,3}(\lambda\mathbf{x})=\frac{\lambda^p\ln^3\lambda}{6}f_{p,0}(\mathbf{x})+\frac{\lambda^p\ln^2\lambda}{2}f_{p,1}(\mathbf{x})+\lambda^p\ln\lambda f_{p,2}(\mathbf{x})+\lambda^p f_{p,3}(\mathbf{x}),\ldots\ . \end{aligned} \tag{6}$$

[6] No other solutions are possible. Substituting $t=x_1$, $t_2=x_2/x_1$, $t_3=x_3/x_1$, ..., $t_n=x_n/x_1$, after differentiating equations (4) with respect to the parameter λ at the point $\lambda = 1$, a system of ordinary differential equations with an independent variable t is obtained. The general solution consists of linear combinations of functions (5), where the constants h_m defining the initial conditions are arbitrary functions of t_2, t_3, ..., t_n. The operations performed are not invertible and therefore it is necessary to verify that functions (5) really satisfy relations (4); however, the verification shows that this condition is not fulfilled for $k > 1$.

Notably, the multiplier $1/k!$ in Eq. (3) provides the equality of coefficients for the lower triangular matrix of linear relations (6):

$$a_{k,j}(\lambda) = \lambda^p \ln^{k-j}(\lambda) / (k-j)!$$

along the diagonals $k - j = const$, as is the case for binomial relations (4).

It can be proved that nontrivial solutions are generally only possible for binomial systems of functional relations of form (4) if $k = 0$ and $k = 1$ (see [31]). However, the fact that [1, 2] assigns binomial functional relations (4) rather than functional relations (6) given as a lower triangular matrix for associated homogeneous functions (3) should be considered an unfortunate inaccuracy, which does not affect subsequent results obtained by the authors of [1, 2].

Indeed, to differentiate relations (4) with respect to parameter p, it is important to take into account that not only the functions $f_{p,k}$ but also the functional coefficients $a(\lambda) = \lambda^p$ and $b(\lambda) = \lambda^p \ln \lambda$ depend on the parameter p. However, as noted earlier, relations (4) are not used in [1, 2] for construction of new generalized associated functions, only playing the role of a formal definition, which was not elaborated further.

Thus, we believe that the authors of numerous subsequent publications using form (4) as a starting point in the formulation and proof of new theorems are not quite correct.

Unfortunately, this truncated definition of associated homogeneous functions has persisted in various publications for over half a century (see, for example, [22–30]). For this reason, the corresponding theorems have much less generality than their authors intended, and, as already noted above, they make sense only for degenerate cases $k = 0$ and $k = 1$ if careful revision is not undertaken both for definitions and for proofs of the corresponding theorems.

It is possible that the problems with the definition of associated homogeneous functions in form (4) are well known to the scholars specializing in them (according to some experts), although we have unable to uncover discussions of this issue in the available literature. However, the rather convincing assumption of these experts about the obviousness and even triviality of both the problem itself and the approach to rectifying it, unfortunately, contradicts the state of affairs that form (4) has long been used in numerous publications and even in authoritative reference books (see, for example, the handbook edited by Kreyn [34], republished abroad).

Apparently, any mathematicians noticing the inconsistency of definitions in [1, 2] will correct it based on their own subjective ideas and preferences, which other researchers may not always agree with. For example, an alternative approach presented in [31] and continued by Albeverio et al. [32, 33] (without excessive emphasis on the problem) as well as similar results in monograph [35] probably were not sufficiently appreciated by the mathematical community, since studies with binomial relations (4) as a starting premise remain popular (see, for example, monograph [30] published in 2012). There is also a risk that instead of a single line of research, a number of disparate definitions will arise for mathematical objects that are close, but do not completely coincide, for which the conditions of equivalence to each other will have to be specifically formulated and proved.

It is worth noting there are certain doubts about the proposal (introduced in [31]) to preserve the term ‘associated homogeneous functions’ for binomial relations (4), which have meaning only when $k = 0$ and $k = 1$, additionally using the term "quasi-associated homogeneous functions" for mathematical objects with a lower triangular matrix of coefficients. In this case, the binomial Gelfand relations, valid for $k = 0$ and $k = 1$, should become special cases of the proposed quasi-associated homogeneous functions. In our opinion, such an approach to the problem would thoroughly belittle the value of the innovative ideas proposed in [1, 2] and definitely contradicts the fact that the associated homogeneous functions are successfully used in the original publications for any orders of k in explicit form. The reinterpretation of the concept in [1, 2] and in [31] is not so drastic that it would be advisable to introduce a new term (see also the comments in the Appendix).

The goal of this study is to prove that linear functional relations are suitable for to correctly define associated homogeneous functions of several real variables as a separate class of functions. These relations can be used as a substitute for differentiation by degree of homogeneity, which requires a certain accuracy, has not been defined unambiguously and is not always feasible.

A number of theorems are proved to achieve this goal, which should demonstrate that this approach is quite productive and deserves further development.

Elementary definitions and theorems

Definition 1. A finite or infinite chain of functions $f_k(\mathbf{x})$ $(k=0,1,2,\ldots)$ is called associated homogeneous functions of the general form of degree p and order k if there exist such functions $a_{k,j}(\lambda)$ and $a(\lambda)\not\equiv 0$ forming a lower triangular matrix of coefficients with the same coefficients $a(\lambda)$ along the main diagonal that the following linear functional relations are satisfied for $\forall\lambda>0$ and $\forall\mathbf{x}\in R^n$:

$$\begin{aligned}&\text{for } k=0,\ f_0(\lambda\mathbf{x})=a(\lambda)f_0(\mathbf{x}),\\&\text{for } k\geq 1,\ f_k(\lambda\mathbf{x})=\sum_{j=0}^{k-1}a_{k,j}(\lambda)f_j(\mathbf{x})+a(\lambda)f_k(\mathbf{x}).\end{aligned}\tag{7}$$

Definition 2. A finite or infinite chain of functions $f_k(\mathbf{x})$ $(k=0,1,2,\ldots)$ is called normalized associated homogeneous functions of degree p and order k if relations (7) are satisfied for $\forall\lambda>0$ and $\forall\mathbf{x}\in R^n$, where the coefficients $a_{k,j}(\lambda)=a_{k-j}(\lambda)$ and $a(\lambda)=a_0(\lambda)\not\equiv 0$ form a lower triangular matrix with the same coefficients along any diagonals $k-j=const$.

Definition 3. A finite or infinite chain of functions $f_k(\mathbf{x})$ $(k=0,1,2,\ldots)$ is called fundamental associated homogeneous functions of degree p and order k if relations (7) are satisfied for $\forall\lambda>0$ and $\forall\mathbf{x}\in R^n$, where the coefficients are given by explicit formulas

$$a_{k,j}(\lambda)=a_{k-j}(\lambda)=\lambda^p\ln^{k-j}(\lambda)/(k-j)!,\ a(\lambda)=a_0(\lambda)=\lambda^p.$$

Definition 1 gives the widest possible conditions of type (4), when the solutions are linear combinations composed of functions (3) with constant but not completely arbitrary coefficients (see Theorems 1 and 4 below).

Definition 2 leads to a narrower class of functions, but it allows you to formulate and prove additional statements that are not generally fulfilled. It uses the idea proposed by the authors of [1, 2] about previously unknown functional coefficients that are identical along the diagonals of the matrix. In this case, the solutions are linear combinations composed of functions (3) with a narrower set of constant but not completely arbitrary coefficients.

Definition 3 sets the closest possible constraints for linear functional relations (7), when one of the possible solutions, or even the unique solution (unfortunately, this requirement is unattainable) are functions of form (3).

Separate statements for normalized associated homogeneous functions and fundamental associated homogeneous functions are not given here if they are corollaries of the corresponding Theorems for associated homogeneous functions of a general form.

Remark 1. Functions in Definition 1 apparently represent the most general case of the approach under consideration, when the theory developed still has a manageable form.

Functions in Definition 2 are called normalized due to the presence of normalizing factors $1/k!$ in Eq. (3). These multipliers are necessary so that the coefficients $a_{k,j}(\lambda)$ for the lower triangular matrix are the same along the diagonals $k-j=const$ (compare with the notations in [31–33], where the authors do not use the normalizing multiplier, and the multipliers before the functional coefficients that arise for this reason are ignored).

Functions in Definition 3 are called fundamental, since any associated homogeneous functions of the general form from Definition 1 are linear combinations of functions of form (3), although their coefficients are not completely arbitrary but rather selected by a special technique (see Theorems 1 and 4 below).

Remark 2. Conditions taking the form $\forall\lambda>0$ and $\forall\mathbf{x}\in R^n$ may seem too rigid for some applications (for example, when considering electron and ion-optical systems with Euler-homogeneous electric and magnetic fields [5–21]). These conditions can be replaced by more practical ones. Let $\Omega\subset R^n$ be some region of n-dimensional space, so that the numbers $\lambda_a(\mathbf{x})$ and $\lambda_b(\mathbf{x})$ are defined for $\forall\mathbf{x}\in\Omega$, satisfying the conditions

$$0 < \lambda_a(\mathbf{x}) < 1 < \lambda_b(\mathbf{x}),$$

for which the points $\lambda\mathbf{x} \in \Omega$ at

$$\lambda_a(\mathbf{x}) \le \lambda \le \lambda_b(\mathbf{x}).$$

Then, with minimal changes in definitions as well as formulations and proofs of theorems, the conditions $\forall \mathbf{x} \in R_n$ and $\forall \lambda > 0$ can be replaced with conditions

$$\forall \mathbf{x} \in \Omega \text{ and } \forall \lambda \in \left[\lambda_a(\mathbf{x}), \lambda_b(\mathbf{x})\right].$$

Remark 3. If the function $f_0(\mathbf{x})$ in relations (7) is continuous at least at one point at which its value is not zero, then the function $a(\lambda)$ is nonzero at $\lambda = 1$, continuous at the point $\lambda = 1$ and thus continuous and strictly positive on the entire positive semi-axis $\lambda > 0$. Furthermore,

$$\forall \lambda, \mu > 0: \; a(\lambda\mu) = a(\lambda)a(\mu)$$

(see Theorem 2 below).

In this case, $a(\lambda) = \lambda^p$ is the only possible option (see monograph [36]), while $f_0(\vec{x})$ is a positively homogeneous Euler function of degree p (see §§ 187, 188 in [3]). Everywhere discontinuous functions $f_0(\vec{x})$ form a special class of functions considered separately from positively homogeneous Euler functions, and Theorem 6 is intended for analyzing them. Theorems 1 and 2 are formulated in such a way that they are applicable to both continuous and everywhere discontinuous functions.

Theorem 1. *If* $\forall \lambda, \; \mu > 0$, *the coefficients* $a(\lambda)$ *and* $a_{k,j}(\lambda)$ *in relations (7) satisfy the equalities*

$$\begin{aligned} &a(\lambda\mu) = a(\lambda)a(\mu), \\ &\forall\, k \ge 1, j < k: \; a_{k,j}(\lambda\mu) = \sum_{s=j}^{k} a_{k,s}(\lambda)a_{s,j}(\mu), \end{aligned} \tag{8}$$

where $a(\lambda) \not\equiv 0, \; a_{k,k}(\lambda) = a(\lambda)$.

Then the following statements hold true.

1. System of relations (7) has solutions other than identically zero, which can be represented as

$$f_0(\mathbf{x}) = \begin{cases} a(x_1)h_0\left(\dfrac{x_2}{x_1}, \dfrac{x_2}{x_1}, \ldots, \dfrac{x_n}{x_1}\right), & x_1 > 0, \\ a(|x_1|)g_0\left(\dfrac{x_2}{x_1}, \dfrac{x_2}{x_1}, \ldots, \dfrac{x_n}{x_1}\right), & x_1 < 0, \end{cases} \tag{9}$$

$$f_k(\mathbf{x}) = \begin{cases} a(x_1)h_k\left(\dfrac{x_2}{x_1}, \dfrac{x_2}{x_1}, \ldots, \dfrac{x_n}{x_1}\right) + \displaystyle\sum_{j=0}^{k-1} a_{k,j}(x_1)h_j\left(\dfrac{x_2}{x_1}, \dfrac{x_2}{x_1}, \ldots, \dfrac{x_n}{x_1}\right), & x_1 > 0, \\ a(|x_1|)g_k\left(\dfrac{x_2}{x_1}, \dfrac{x_2}{x_1}, \ldots, \dfrac{x_n}{x_1}\right) + \displaystyle\sum_{j=0}^{k-1} a_{k,j}(|x_1|)g_j\left(\dfrac{x_2}{x_1}, \dfrac{x_2}{x_1}, \ldots, \dfrac{x_n}{x_1}\right), & x_1 < 0, \end{cases} \tag{10}$$

where h_j, g_j are *arbitrary functions of* $(n-1)$ *variables.*

2. System of relations (7) has no other solutions: any solution can be represented as (9), (10) with proper selection of functions h_j, g_j.

3. The following conditions are satisfied for the functions given by Eqs. (9) and (10):

$$\begin{aligned} &h_j(t_2, t_3, \ldots, t_n) = f_j(1, t_2, t_3, \ldots, t_n), \\ &g_j(t_2, t_3, \ldots, t_n) = f_j(-1, -t_2, -t_3, \ldots, -t_n). \end{aligned} \tag{11}$$

Proof. Eq. (9) for separate cases $x_1 > 0$ and $x_1 < 0$ for positively homogeneous functions follows from the arguments used in [3, §§ 187, 188] for homogeneous Euler functions.

We follow the induction method: let parameterization (10) for $x_1 > 0$ hold true for functions f_j at $j = 0,1,\ldots,k-1$. Let us write the function f_k in the form (11), where the functions h_j at $j = 0,1,\ldots,k-1$ are inherited from the previous steps, and an arbitrary function $h_k\left(x_1, x_2/x_1, \ldots, x_n/x_1\right)$ of n variables is used as h_k. Apparently, freedom of choice of the function f_k is not limited in any way at this step. After substituting this expression into relation (7) with subscript k and additional algebraic transformations, we obtain the condition

$$\forall \lambda > 0: \; h_k\left(\lambda x_1, \frac{x_2}{x_1}, \ldots, \frac{x_2}{x_1}\right) = h_k\left(x_1, \frac{x_2}{x_1}, \ldots, \frac{x_2}{x_1}\right).$$

It follows that the function h_k does not depend on the first argument, which means that condition (10) must be fulfilled for the function f_k satisfying the relations (7). All algebraic transformations are invertible, therefore, it follows from the representation of the function f_k in the form (10), where the function h_k does not depend on the first argument, that relations (7) for the subscript k hold true. The reasoning for $x_1 < 0$ is repeated for $x_1 > 0$.

Theorem 1 is proved.

Remark 4. Conditions (8) are sufficient but not necessary for relations (8) to have nontrivial solutions. However, as Theorems 2 and 3 (see below) show, conditions (8) are rather general, and examples when they are not satisfied are rare exceptions to the common rule. For normalized associated homogeneous functions (see Definition 2), conditions (8) are not only sufficient but also necessary if the function f_0 is not identically zero. The functions $a_{k,j}(\lambda)$ and $a(\lambda)$ in Definition 3 themselves satisfy relations (8). In particular, this is the reason why the most common notation for fundamental associated homogeneous functions has the form (9) and (10) after substituting the following functions into these formulas:

$$a_{k,j}(\lambda) = \lambda^p \ln^{k-j}(\lambda)/(k-j)!, \; a(\lambda) = a_{k,k}(\lambda) = \lambda^p.$$

Theorem 2. *If relations (7) have a solution in the form of linearly independent functions $f_k(\mathbf{x})$, then, for $\forall \lambda, \mu > 0$, the coefficients $a(\lambda)$ and $a_{k,j}(\lambda)$ are determined uniquely and must satisfy relations (8).*

Proof. If $f_0(\mathbf{x}) \neq 0$ at least at one point, then, if $k = 0$, conditions (7) lead to equalities of the form

$$f_0(\lambda\mu\mathbf{x}) = a(\lambda\mu) f_0(\mathbf{x}) = a(\lambda) f_0(\mu\mathbf{x}) = a(\lambda) a(\mu) f_0(\mathbf{x}),$$

that is, for $\forall \lambda, \; \mu > 0: a(\lambda\mu) = a(\lambda) a(\mu)$.. It follows, in particular, that $a(\lambda)$ is either identically zero or strictly greater than zero at all points.

By induction, if $k > 0$, the remaining relations (8) are obtained from relations (7) considered at point $\lambda\mu\mathbf{x}$. If we assume that elations (7) are satisfied for a certain set of linearly independent functions $f_k(\mathbf{x})$ simultaneously for two sets of functions $a_{k,j}(\lambda)$ and $\overline{a}_{k,j}(\lambda)$, then the equalities $a(\lambda) \equiv \overline{a}(\lambda)$, $\forall k > j: \; a_{k,j}(\lambda) \equiv \overline{a}_{k,j}(\lambda)$ follow immediately.

Theorem 2 is proved.

Theorem 2. *If $a(\lambda) \neq 0$ in relations (7) and the coefficients $a_{k,k-1}(\lambda)$ for $\forall k \geq 1$ are not identically zero, then any chain of functions $f_k(\mathbf{x})$ satisfying relations (7) consists of linearly independent functions, if the initial function $f_0(\mathbf{x})$ is not identically zero.*

Proof. Let us assume that f_k is the first function that is linearly dependent on the previous functions f_j $(j = 0,1,\ldots,k-1)$. We represent f_k as a linear combination of functions f_j with constant coefficients and substitute it into relation (7) with the subscript k. Let us group the multipliers at different functions f_j. The multiplier for the function f_{k-1} will turn out to be equal to $a_{k,k-1}(\lambda)$ and should be equal to zero due to the linear independence of the functions f_j; however, this leads to a contradiction, provided that $a_{k,k-1}(\lambda)$ is not equal to zero for at least one value of λ.

Theorem 3 is proved.

Theorem 4. *If each of the coefficients* $a(\lambda)$ *and* $a_{k,j}(\lambda)$ *in relations (8) is continuous at least at one point, and* $a(\lambda) \not\equiv 0$*, then the following statements hold true.*

1. These coefficients take the form

$$a(\lambda) = \lambda^p \text{ and } a_{k,j}(\lambda) = \lambda^p b_{k,j}(\ln\lambda),$$

where $b_{k,j}(t)$ *are polynomials of degree* $k - j$*, which ensures differentiability of any order for functions* $a(\lambda)$ *and* $a_{k,j}(\lambda)$.

2. Polynomials $b_{k,j}$ *are uniquely determined from the following differential recurrence relations:*

$$\begin{aligned} &\text{for } \forall k \geq 0,\ b_{k,k}(t) \equiv 1, \\ &\text{for } \forall k \geq 1,\ j = k-1, k-2, \ldots, 0\ \ b'_{k,j}(t) = \sum_{s=j}^{k-1} p_{k,s} b_{s,j}(t);\ b_{k,j}(0) = 0, \end{aligned} \tag{12}$$

or (which produces the same result) of the following differential recurrence relations:

$$\begin{aligned} &\text{for } \forall k \geq 0,\ b_{k,k}(t) \equiv 1, \\ &\text{for } \forall k \geq 1, j = k-1,\ k-2, \ldots, 0,\ b'_{k,j}(t) = \sum_{s=j+1}^{k} b_{k,s}(t)\, p_{s,j};\ b_{k,j}(0) = 0, \end{aligned} \tag{13}$$

where $p_{k,j}$ *are arbitrary constants, and* $a'_{k,j}(1) = b'_{k,j}(0) = p_{k,j}$.

3. Equivalent recurrence conditions (12) and (13) are both necessary and sufficient to fulfill relations (8) with the functions $a(\lambda)$ *and* $a_{k,j}(\lambda)$.

Proof. The condition $a(\lambda\mu) = a(\lambda)a(\mu)$ for functions $a(\lambda)$ that are nonzero and continuous at least at one point implies the condition $a(\lambda) = \lambda^p$ with some real exponent p. After adopting, instead of functions $a_{k,j}$, new functions $b_{k,j}$ in accordance with the equalities:

$$\begin{aligned} &\text{for } j < k\ \ a_{k,j}(\lambda) = \lambda^p b_{k,j}(\ln\lambda), \\ &\text{for } j = k\ \ a(\lambda) = a_{k,k}(\lambda) = \lambda^p b_{k,k}(\ln\lambda) = \lambda^p, \end{aligned}$$

equivalent additive relations for functions $b_{k,j}$ are obtained from multiplicative relations (8) after substitution $y = \ln\lambda$, $z = \ln\mu$:

$$\begin{aligned} &\text{for } \forall k \geq 1, 0 \leq j < k\ \ b_{k,j}(y+z) = \sum_{s=j}^{k} b_{k,s}(y) b_{s,j}(z), \\ &\text{for } \forall k \geq 0,\ b_{k,k}(y) = 1, \end{aligned} \tag{14}$$

where the following initial conditions are also satisfied for the functions $b_{k,k}$ and $b_{k,j}$:

$$\begin{aligned} &\text{for } \forall k \geq 1,\ 0 \leq j < k\ \ b_{k,j}(0) = 0, \\ &\text{for } \forall k \geq 0,\ b_{k,k}(0) = 1 \end{aligned} \tag{15}$$

(equalities (15) are derived by induction from the condition $b_{k,k}(y) \equiv 1$ and relations (14) considered with $y = z = 0$).

Relations (12) and (13) are necessary conditions for differentiable functions: relations (12) are obtained by differentiating relations (14) with respect to y at point $y = 0$ and substituting $z \to t$, and relations (13) are obtained by differentiation with respect to z at point $z = 0$ and substituting $y \to t$. Using a system of conditions (12) or a system of conditions (13), functions (polynomials) $b_{k,j}$ with sequentially iterated subscripts $k = 0, 1, 2, \ldots$, $j = k, k-1, \ldots 0$ are reconstructed uniquely as soon as a set of constants $p_{k,j}$ ($k > 0$, $0 \leq j \leq k-1$) is found.

To reverse the transition from relations (12) to relations (14), consider the following functions for $k \geq 0$ and $0 \leq j \leq k$:

$$\Phi_{k,j}(y,z) = b_{k,j}(y+z) - \sum_{s=j}^{k} b_{k,s}(y) b_{s,j}(z).$$

If relations (12) are satisfied, the derivative of the function $\Phi_{k,j}(y,z)$ with respect to the variable y for $j < k$, by virtue of the condition $b'_{k,k}(y) \equiv 0$, obeys the equality

$$\frac{\partial \Phi_{k,j}(y,z)}{\partial y} = b'_{k,j}(y+z) - \sum_{m=j}^{k} b'_{k,m}(y) b_{m,j}(z) =$$

$$= \sum_{s=j}^{k-1} p_{k,s} b_{s,j}(y+z) - \sum_{m=j}^{k-1} \left(\sum_{s=m}^{k-1} p_{k,s} b_{s,m}(y) b_{m,j}(z) \right) =$$

$$= \sum_{s=j}^{k-1} p_{k,s} b_{s,j}(y+z) - \sum_{s=j}^{k-1} \left(\sum_{m=j}^{s} p_{k,s} b_{s,m}(y) b_{m,j}(z) \right) =$$

$$= \sum_{s=j}^{k-1} p_{k,s} \left(b_{s,j}(y+z) - \sum_{m=j}^{s} b_{s,m}(y) b_{m,j}(z) \right) =$$

$$= \sum_{s=j}^{k-1} p_{k,s} \Phi_{s,j}(y,z).$$

If $\partial \Phi_{k,j}(y,z) / \partial y \equiv 0$, then $\Phi_{k,j}(y,z) \equiv \Phi_{k,j}(0,z) \equiv 0$. In addition, $\Phi_{k,k}(y,z) \equiv 0$ for $\forall k \geq 0$. Therefore, by induction, $\Phi_{k,j}(y,z) \equiv 0$ under any $k \geq 0$, $0 \leq j \leq k$. This means that relations (14) hold true if relations (12) are satisfied.

Similarly, the sufficiency of relations (13) for relations (14) is proved by differentiating the function $\Phi_{k,j}(y,z)$ with respect to z. Since conditions (12) follow conditions (14), and conditions (14) follow conditions (13) and vice versa, systems of recurrence relations (12) and (13) are equivalent to each other (which is not always immediately obvious).

Now let the functions $b_{k,j}$ be merely continuous at individual points. It is required to prove that in this case the functions $b_{k,j}(t)$, if they obey conditions (14), are differentiable at all points and, as a result, are determined by recurrence relations (12) or (13).

By definition, $b_{k,k}(t) = 1$, and for $k \geq 1$, $j = k-1$, relation (8) is reduced to the Cauchy additive functional equation with respect to the function $b_{k,k-1}$. Its solution, due to its continuity at least at one point, must be a linear and thus everywhere differentiable function: $b_{k,k-1}(t) = p_{k,k-1} t$ [36]. Such a function $b_{k,k-1}(t)$ obeys relations (12) and (13) with a constant $p_{k,k-1}$.

We follow the induction method: let $0 \leq j \leq k-2$, then the required statement is proved for the functions $b_{k,k}(t)$, $b_{k,k-1}(t)$, ..., $b_{k,j+1}(t)$. Let us use recurrence relations (12) (the reasoning is similar for relations (13)).

We construct the function $\tilde{b}_{k,j}(t)$ using the condition

$$\tilde{b}'_{k,j}(t) = \sum_{s=j+1}^{k-1} b_{k,s}(t) p_{s,j}; \tilde{b}_{k,j}(0) = 0.$$

It corresponds to condition (12) with a constant $p_{k,j} = 0$ and constants $p_{s,j}$ inherited from previous induction steps. Such a function $b_{k,j}(t)$ is infinitely differentiable and satisfies condition (14):

$$\tilde{b}_{k,j}(y+z) = \tilde{b}_{k,j}(y) b_{j,j}(z) + \sum_{s=j+1}^{k-1} b_{k,s}(y) b_{s,j}(z) + b_{k,k}(y) \tilde{b}_{k,j}(z) =$$

$$= \tilde{b}_{k,j}(y) + \sum_{s=j+1}^{k-1} b_{k,s}(y) b_{s,j}(z) + \tilde{b}_{k,j}(z).$$

Let $b_{k,j}(t) = \tilde{b}_{k,j}(t) + c(t)$, where the auxiliary function $c(t)$ must be continuous at the same points as the function $b_{k,j}(t)$. Condition (14) is satisfied for the function $b_{k,j}(t)$. This means that the function $c(t)$ satisfies the Cauchy additive equation and, due to its continuity, it is a linear function differentiable everywhere: $c(t) = p_{k,j}t$ [36]. Therefore, the function $b_{k,j}(t)$ is infinitely differentiable at all points and satisfies relations (12) and (13), where a constant $p_{k,j}$ is added to the list of constants.

Theorem 4 is proved.

Remark 5. The polynomials $b_{k,j}(t)$ are determined unambiguously from the differential recurrence relations (12) or (13), as soon as the constants $p_{k,j}$ are found. For normalized associated homogeneous functions, the equalities $b_{k,j}(t) = c_{k-j}(t)$ and $p_{k,j} = q_{k-j}$ hold true, where the polynomials $c_k(t)$ are determined from the recurrence relations

$$c_0(t) = 1;\ \forall k \geq 1:\ c_k'(t) = \sum_{s=1}^{k} q_s c_{k-s}(t) = \sum_{s=0}^{k-1} q_{k-s} c_s(t),\ c_k(0) = 0$$

including the constants q_k. Finally, for fundamental associated homogeneous functions $p_{k,j} = 0$ at $j = 0,1,...,k-2$, $p_{k,k-1} = 1$ and $b_{k,j}(t) = t^{k-j}/(k-j)!$

Theorem 5 (Euler criterion). *Let the chain of functions $f_k(\mathbf{x})$ $(k = 0,1,2,\ldots)$ consist of differentiable functions for which, with $\forall \mathbf{x} \in R^n$ (possibly excluding the point $\mathbf{x} = 0$), the following relations are satisfied:*

$$\begin{aligned} &\text{for } k = 0,\ x_1 \frac{\partial f_0(\mathbf{x})}{\partial x_1} + x_2 \frac{\partial f_0(\mathbf{x})}{\partial x_2} + \cdots + x_n \frac{\partial f_0(\mathbf{x})}{\partial x_n} = p f_0(\mathbf{x}), \\ &\text{for } k \geq 1,\ x_1 \frac{\partial f_k(\mathbf{x})}{\partial x_1} + x_2 \frac{\partial f_k(\mathbf{x})}{\partial x_2} + \cdots + x_n \frac{\partial f_k(\mathbf{x})}{\partial x_n} = p f_k(\mathbf{x}) + \sum_{j=0}^{k-1} p_{k,j} f_j(\mathbf{x}) \end{aligned} \tag{16}$$

with some constants p and $p_{k,j}$.

Then the following statements hold true.

Functions $f_k(\mathbf{x})$ satisfy relations (7), where

$$a(\lambda) = \lambda^p \text{ and } a_{k,j}(\lambda) = \lambda^p b_{k,j}(\ln \lambda),$$

where the functions $b_{k,j}(y)$ are *polynomials of degree $k - j$ satisfying recurrence relations (12) and (13) with the constants $p_{k,j}$ included in relations (16).*

2. If the functions $f_k(\mathbf{x})$ differentiable at $\forall \mathbf{x} \in R^n$ (possibly with the exception of point $\mathbf{x} = 0$) and the functions $a(\lambda)$ and $a_{k,j}(\lambda)$ differentiable at $\forall \lambda \geq 0$ (possibly with the exception of $\lambda = 0$) satisfy the system of functional relations (8), then the relations (16) with the constants $p = a'(1)$ and $p_{k,j} = a_{k,j}'(1)$ are satisfied.

Proof. Like Euler's differential criterion for homogeneous functions (see §§ 187, 188 in [3]), differential relations (16) necessarily follow from relations (7) after their differentiation with respect to λ at the point $\lambda = 1$. To prove the sufficiency of relations (16), by analogy with the calculations in the above-mentioned paragraphs of [3], the functions $\Omega_k(\lambda, \mathbf{x})$ are used:

$$\Omega_k(\lambda, \mathbf{x}) = \sum_{j=0}^{k} \lambda^{-p} b_{k,j}(-\ln \lambda) f_j(\lambda \mathbf{x}),$$

where the functions $b_{k,j}(t)$ are calculated in accordance with relations (12) and (13) including constants $p_{k,j}$ from relations (16).

When differentiating functions $\Omega_k(\lambda, \mathbf{x})$ with respect to λ, a weighted sum of relations (16), calculated at the point $\lambda \mathbf{x}$, is obtained at points $\mathbf{x} \neq 0$, which are zero for all λ and $\mathbf{x}$. Since the derivative of the function $\Omega_k(\lambda, \mathbf{x})$ with respect to the parameter λ is identically zero, the identity $\Omega_k(\lambda, \mathbf{x}) \equiv \Omega_k(1, \mathbf{x})$ is fulfilled; so after substitutions $\lambda \to 1/\mu$, $\mathbf{x} \to \mu \mathbf{x}$, $\mu \to \lambda$, relations (7) for the subscript k with the functions $a(\lambda) = \lambda^p$ and $a_{k,j}(\lambda) = \lambda^p b_{k,j}(\ln \lambda)$ are obtained.

Theorem 5 is proved.

Remark 6. Theorem 5 for associated homogeneous functions of a general form is an equivalent of Euler's criterion for homogeneous functions (see §§ 187, 188 in [3]). For normalized associated functions, where $a_{k,j}(t) = a_{k-j}(t)$ (see Definition 2), equality $p_{k,j} = q_{k-j}$ holds true in relations (16), where the constants q_j are determined by conditions $q_j = a_j(1)$.

Therefore, the Euler criterion for normalized associated homogeneous functions has the form

$$\text{for } k = 0,\ x_1 \frac{\partial f_0(\mathbf{x})}{\partial x_1} + x_2 \frac{\partial f_0(\mathbf{x})}{\partial x_2} + \cdots + x_n \frac{\partial f_0(\mathbf{x})}{\partial x_n} = p f_0(\mathbf{x}),$$

$$\text{for } k \geq 1,\ x_1 \frac{\partial f_k(\mathbf{x})}{\partial x_1} + x_2 \frac{\partial f_k(\mathbf{x})}{\partial x_2} + \cdots + x_n \frac{\partial f_k(\vec{x})}{\partial x_n} = p f_k(\mathbf{x}) + \sum_{j=1}^{k} q_j f_{k-j}(\mathbf{x}).$$

Accordingly, for fundamental associated functions (see Definition 3) the equalities $p_{k,k-1} = 1$ and $p_{k,j} = 0$ hold true for $0 \leq j \leq k-2$ in relations (16), and the Euler criterion takes the following form:

$$\text{for } k = 0,\ x_1 \frac{\partial f_0(\mathbf{x})}{\partial x_1} + x_2 \frac{\partial f_0(\mathbf{x})}{\partial x_2} + \cdots + x_n \frac{\partial f_0(\mathbf{x})}{\partial x_n} = p f_0(\mathbf{x}),$$

$$\text{for } k \geq 1,\ x_1 \frac{\partial f_k(\mathbf{x})}{\partial x_1} + x_2 \frac{\partial f_k(\mathbf{x})}{\partial x_2} + \cdots + x_n \frac{\partial f_k(\mathbf{x})}{\partial x_n} = p f_k(\mathbf{x}) + f_{k-1}(\mathbf{x}).$$

Remark 7. There are examples when the same chain of differentiable functions $f_k(\mathbf{x})$ simultaneously satisfies several systems of form (7) with different coefficients $a_{k,j}(\lambda)$. In this case, the functions

$$a(\lambda) = \lambda^p \text{ and } a_{k,j}(\lambda) = \lambda^p b_{k,j}(\ln \lambda)$$

correspond to only one of the many possible systems of form (7) for the functions $f_k(\mathbf{x})$ under consideration. Here $b_{k,j}(y)$ are determined from recurrence relations (12) or (13) with the constants $p_{k,j}$ taken from equalities (16).

However, if the functions $f_k(\mathbf{x})$ are linearly independent (for which it is sufficient that for $\forall k \geq 1$: $p_{k,k-1} \neq 0$ and that the function $f_0(\mathbf{x})$ is not identical to zero (see Theorem 3), then the coefficients $a(\lambda)$ and $a_{k,j}(\lambda)$ of system (7) are defined uniquely (see Theorem 2) and thus must coincide with the expressions

$$a(\lambda) \text{ and } a_{k,j}(\lambda) = \lambda^p b_{k,j}(\ln \lambda),$$

used in the proof of Theorem 5.

Associated homogeneous functions discontinuous at all points

If each of the coefficients $a(\lambda)$ and $a_{k,j}(\lambda)$ in relations (7) is continuous at least at one point, then Theorem 4 gives an exhaustive answer to the question of what form the coefficients $a(\lambda)$ and $a_{k,j}(\lambda)$ should have in order for functional relations (8) to be satisfied and for system (7) to have nontrivial solutions (the general form of these solutions is given in Theorem 1). In this case, the coefficients $a(\lambda)$ and $a_{k,j}(\lambda)$, continuous at least at one point, will be infinitely differentiable at all points $\lambda > 0$. However, there are everywhere discontinuous coefficients $a(\lambda)$ and $a_{k,j}(\lambda)$, for which relations (8) are satisfied, and thus, in accordance with Theorem 1, there are nontrivial everywhere discontinuous solutions of system (7).

Theorem 6. *Let* $\theta_0(t)$ *and* $\theta_{k,j}(t)$ *be arbitrary (generally speaking, everywhere discontinuous) solutions of the Cauchy additive functional equation [36]:*

$$\theta(x+y) = \theta(x) + \theta(y).$$

Then the general solution for functional relations (8) are the functions

$$a(\lambda)=\exp\left(\theta_0\left(\ln\lambda\right)\right) \text{ and } a_{k,j}(\lambda)=\exp\left(\theta_0\left(\ln\lambda\right)\right)\varphi_{k,j}\left(\ln\lambda\right),$$

where functions $\varphi_{k,j}(t)$ *are obtained through substitutions* $p_{k,j}\to\theta_{k,j}(t)$, $y\to 1$ *from functions* $b_{k,j}(y)$ *that are given by the recurrence relations (12) or (13). At* $\theta_0(t)=pt$ *and* $\theta_{k,j}(t)=p_{k,j}t$, *these solutions coincide with the functions used in Theorem 4.*

Unfortunately, we do not have a sufficiently short proof of Theorem 6 at our disposal to present it in this paper at least concisely. This goes beyond the scope of this study and requires separate consideration. A separate publication will be dedicated to the proof of this theorem.

Mutually homogeneous associated functions

The functions considered in [39] correspond to the case when the fundamental associated homogeneous functions are complex-valued and can be represented as the sum of real and imaginary parts that are functionally independent of each other, and the homogeneity degree is a complex number. Then each of Eqs. (7) turns into a pair of linear functional relations of the same length, written for pairs of mutually homogeneous associated functions. Similarly, Eqs. (7) of the general form will also be transformed by adopting complex-valued functions and complex homogeneity degrees.

Apparently, it will not be possible to obtain any other classes of functions from such paired relations, written in the most general form, but this problem requires separate study. It also remains an open question whether it is possible to obtain interesting mathematical objects if triplets, quadruplets, etc., of functions characterized by dense matrices of linear functional relations of a general form are similarly combined into a single whole [37], with the resulting equations then extended to an infinite block-triangular structure of associated linear functional relations. The particular case when the coefficient matrix is block-triangular (instead of the lower triangular matrix corresponding to linear functional relations (7)) and the «seed» of the chain of associated functions consists of several homogeneous functions with different degrees of homogeneity is briefly discussed later in the section “Vector-associated homogeneous functions”.

Generalized associated homogeneous functions

If we apply the approach proposed in [1, 2], then the results obtained in the previous section can be extended to generalized functions (linear continuous functionals defined in the corresponding linear space of trial functions). This allows to correctly define generalized associated homogeneous functions of a general form, generalized normalized associated homogeneous functions and generalized fundamental homogeneous functions.

Evidently, everywhere discontinuous (and thus nowhere integrable) associated homogeneous functions cannot have equivalents in the class of generalized functions. A more detailed consideration of these interesting problems (in particular, the formulation and proof of Euler’s differential criterion for generalized associated homogeneous functions) is planned to be presented in a separate specialized paper.

Vector-associated homogeneous functions

A further development of Gelfand’s idea of associated homogeneous functions are vector-associated functions, where the initial «seed» for a chain of functions are several Euler-homogeneous functions with different degrees of homogeneity.

Definition 4. A chain of functions $f_k^{(i)}(\vec{x})$ ($k=0,1,2,\ldots$; $i=1,2,\ldots m$) is called m-associated (vector-associated) homogeneous functions of a general form if the following conditions are satisfied for $\forall\lambda>0$ and $\forall\vec{x}\in R^n$ with $i=1,2,\ldots m$:

$$\begin{aligned}&\text{for } k=0,\ f_0^{(i)}(\lambda\mathbf{x})=a_i(\lambda)f_0^{(i)}(\mathbf{x}),\\&\text{for } k\geq 1,\ f_k^{(i)}(\lambda\mathbf{x})=\sum_{j=0}^{k-1}\sum_{s=1}^{m}a_{k,j}^{(i,s)}(\lambda)f_j^{(s)}(\mathbf{x})+a_i(\lambda)f_k^{(i)}(\mathbf{x}),\end{aligned}\tag{17}$$

with some functions that are not pre-defined:

$$a_i(\lambda) \text{ and } a_{k,j}^{(i,s)}(\lambda) \ (k=1,2,\ldots;\ j=0,1,\ldots,k-1;\ i=1,2,\ldots m;\ s=1,2,\ldots m),$$

where for $\forall\lambda>0$, $a_i(\lambda)\neq 0$ and $a_i(\lambda)\neq a_j(\lambda)$ for $i\neq j$.

Remark 8. If each of the functions $a_i(\lambda)$ is continuous at least at one point, then $a_i(\lambda)=\lambda^{p_i}$ and $f_k^{(i)}(\mathbf{x})$ are positively homogeneous functions (see §§ 187, 188 in [3]) with different degrees of homogeneity p_i, satisfying the condition $p_i\neq p_s$ for $i\neq s$.

For vector-associated homogeneous functions, the equivalents of the statements given above hold true (for the most part). In particular, the concept of vector-associated homogeneous functions can be transferred without any loss to the class of generalized functions.

Neither the formulations of these equivalents nor their proofs are given in this study, since their consideration goes beyond the scope of the main topic of the paper.

Regrettably, this far, we have been unable to answer the question about the general form of everywhere discontinuous solutions for relations (17).

Conclusion

The paper continues and summarizes the research presented earlier in [37–39], considering the basic concepts related to the general theory of functions of a real variable. It is assumed that associated homogeneous Gelfand functions can be effective for generalizing Golikov's principle of similarity for trajectories in Eulerian fields to a wider class of electrostatic potentials if we focus on the similarity of paraxial trajectories.

The theorems formulated and proved in this paper are new original results that we obtained for the first time, previously not publicly available.

It is proved that it is possible to construct a detailed consistent theory of associated homogeneous functions of real variables determined using a chain of linear recurrent functional relations of a general type, and also that this approach is quite productive and deserves further development.

The proofs of individual theorems in the paper are sometimes given here in a concise form, since they are not as practically oriented as the formulations of the theorems themselves. It is assumed that in most cases complete proof can be reconstructed in all details by readers who are mathematics graduates, whereas the details and aspects of the proofs that require special attention are of interest mainly to dedicated specialists. The publication of these materials (if continued by the authors) will probably be presented in the relevant specialized journals.

Appendix

Definition of a new class of functions in the original article

In this section, we briefly explore the logic behind defining a new class of functions (associated homogeneous Gelfand–Shapiro functions, ordinary and generalized), described in [1, 2], as we understand it. Additional comments are provided for the aspects that we deem crucial for the discussion.

Thus, the fragments of the Appendix marked as 'Statements' represent a brief retelling of the corresponding sections of the original study [1] (with the utmost respect for this classical work), whereas the fragments of the Appendix marked as 'Comments' have been entirely constructed by the authors of this paper.

Statement 1A. For a linear operator $Af(\mathbf{x})=f(\lambda\mathbf{x})$ giving a similar transformation of function arguments, homogeneous Euler functions of degree p that satisfy the identity $f(\lambda\mathbf{x})=\lambda^p f(\mathbf{x})$ are eigenfunctions with eigenvalues λ^p. In a typical case, along with the eigenfunction f_0 corresponding to the given eigenvalue α, a linear transformation (operator) also has associated eigenfunctions f_1, f_2, ... f_k, ... of various orders that satisfy the relations

$$\begin{aligned} Af_0&=\alpha f_0,\\ Af_1&=\alpha f_1+\beta f_0,\\ Af_2&=\alpha f_2+\beta f_1,\\ &\cdots\\ Af_k&=\alpha f_k+\beta f_{k-1},\\ &\cdots \end{aligned} \tag{1A}$$

where k is the order of the function; α is the eigenvalue under consideration; β is some constant that, with proper normalization of the associated eigenfunctions, can be taken to equal unity (in the general case $\beta \neq 0$).

Therefore, for homogeneous Euler functions, we can try to determine the equivalent of the associated eigenfunctions, i.e., the attached homogeneous functions.

Comment 1A. Strictly speaking, the lower triangular matrix of linear relations with the same coefficients α along the main diagonal can be used with equal success to determine the associated eigenfunctions of linear operators. Indeed, in their physical meaning, associated eigenfunctions are such chains of linearly independent functions for which the result of the action of a given linear operator will be a linear combination composed of the function itself (with a multiplier equal to its eigenvalue) and the preceding functions. At the same time, for a fixed eigenvalue α, it is always possible to move from the lower triangular matrix of equalities

$$Af_k = \sum_{j=0}^{k-1} \beta_{k,j} f_j + \alpha f_k \text{ (where } \beta_{k,k-1} \neq 0)$$

to binomial relations (1A) with normalized coefficients $\beta = 1$ along the diagonal $k - j = 1$, a linear combination of the preceding functions f_1, f_2, ..., f_{k-1} with the corresponding weight coefficients is properly subtracted from each function f_k for $\forall k \geq 2$. Similarly, we can always adopt, instead of the binomial relations (1A) (which, of course, are much easier to work with), an almost arbitrary lower triangular matrix, for which the condition $\beta_{k,k-1} \neq 0$ is the only constraint.

However, homogeneous Euler functions, i.e., eigenfunctions that do not contain a parameter but belong to a single-parameter family of linear similarity operators with a single-parameter set of eigenvalues, are a special case. In view of this, it is necessary to consider associated eigenfunctions that do not contain a parameter and are characterized by single-parameter linear associated relations (if such functions exist).

Apparently, the transition from a single-parameter lower triangular matrix to a single-parameter two-diagonal one through linear replacement of functions that do not contain a parameter cannot be performed in this case. The two methods cease to be equivalent, and we have to deal with a full-fledged lower triangular coefficient matrix, which is not so convenient. The meaningful difference between binomial relations and relations in the form of a lower triangular matrix is not so fundamental that there is a real need to introduce a new term in addition to the Gelfand term, as done in [31].

Statement 2A. A first-order associated homogeneous function of degree p is a function f_1 of real variables, which for any $\lambda > 0$ satisfies the equality

$$f_1(\lambda \mathbf{x}) = \lambda^p f_1(\mathbf{x}) + h(\lambda) f_0(\mathbf{x}), \tag{2A}$$

where $f_0(\mathbf{x})$ is a homogeneous function of real variables of degree p.

It is proved that a continuous function $h(\lambda)$ must have the form

$$h(\lambda) = c\lambda^p \ln \lambda,$$

where $c \neq 0$ is an arbitrary constant.

Due to normalization of the function f_1, it can be assumed that $c = 1$ without loss of generality.

As an example of a first-order associated homogeneous function with zero degree of homogeneity, consider the function $\ln|\mathbf{x}|$ that for $\lambda > 0$ satisfies the condition

$$\ln|\lambda \mathbf{x}| = \ln|\mathbf{x}| + \ln \lambda$$

(unity is a homogeneous function of zero degree).

Statement 3A. A generalized homogeneous function of degree p is defined as a linear continuous functional T_p given over a linear space of infinitely differentiable functions $\varphi(\mathbf{x})$ of real variables of dimension n tending to zero at $|\mathbf{x}| \to \infty$ faster than any power function $1/|\mathbf{x}|^k$ $(k > 0)$ if this functional obeys the identity

$$T_p\left[\varphi\left(\frac{\mathbf{x}}{\lambda}\right)\right]=\lambda^{p+n}T_p\left[\varphi(\mathbf{x})\right]. \tag{3A}$$

Condition (3A) is constructed by analogy with the relations

$$T_p\left[\varphi(\mathbf{x})\right]=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(\mathbf{x})\varphi(\mathbf{x})\,dx_1\ldots dx_n,$$

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(\lambda\mathbf{x})\varphi(\mathbf{x})\,dx_1\ldots dx_n=\lambda^{-n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(\mathbf{y})\varphi\left(\frac{\mathbf{y}}{\lambda}\right)dy_1\ldots dy_n=$$

$$=\lambda^{p}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(\mathbf{x})\varphi(\mathbf{x})\,dx_1\ldots dx_n,$$

which are satisfied when the prototype of the generalized function is a homogeneous function $f(\mathbf{x})$ of real variables, characterized by the degree of homogeneity p.

As an example, it is proved that the n-dimensional delta function is a generalized homogeneous function of degree $-n$ in the sense of definition (3A).

Similarly, by analogy with equality (2A) at $h(\lambda)=\lambda^p\ln\lambda$, which must be satisfied for ordinary first-order associated homogeneous functions, a generalized first-order associated function $T_{p,1}$ defined as a linear continuous functional that obeys the identity

$$T_{p,1}\left[\varphi\left(\frac{\mathbf{x}}{\lambda}\right)\right]=\lambda^{p+n}T_{p,1}\left[\varphi(\mathbf{x})\right]+\lambda^{p+n}\ln\lambda\,T_{p,0}\left[\varphi(\mathbf{x})\right], \tag{4A}$$

where $T_{p,0}$ is some generalized homogeneous function of degree p, not equal to zero.

Finally, by analogy with the relations (1A), a chain of linear continuous functionals $T_{p,k}$ is called generalized associated functions of order k and degree p if the following identities are satisfied:

$$T_{p,k}\left[\varphi\left(\frac{\mathbf{x}}{\lambda}\right)\right]=\lambda^{p+n}T_{p,k}\left[\varphi(\mathbf{x})\right]+\lambda^{p+n}\ln\lambda\cdot T_{p,k-1}\left[\varphi(\mathbf{x})\right]. \tag{5A}$$

Comment 2A. There is an inaccuracy in the reasoning here, since (as established, in particular, in [31]) there are no functionals $T_{p,k}$ with such properties even for $k\geq 2$. It follows from the chain of equalities

$$T_{p,k+2}\left[\varphi\left(\frac{\vec{x}}{\lambda\mu}\right)\right]=\lambda^{p+n}T_{p,k+2}\left[\varphi\left(\frac{\mathbf{x}}{\mu}\right)\right]+\lambda^{p+n}\ln\lambda\,T_{p,k+1}\left[\varphi\left(\frac{\mathbf{x}}{\mu}\right)\right]=$$

$$=\lambda^{p+n}\left(\mu^{p+n}T_{p,k+2}\left[\varphi(\mathbf{x})\right]+\mu^{p+n}\ln\mu\,T_{p,k+1}\left[\varphi(\mathbf{x})\right]\right)+$$

$$+\lambda^{p+n}\ln\lambda\left(\mu^{p+n}T_{p,k+1}\left[\varphi(\mathbf{x})\right]+\mu^{p+n}\ln\mu\,T_{p,k}\left[\varphi(\mathbf{x})\right]\right),$$

$$T_{p,k+2}\left[\varphi\left(\frac{\mathbf{x}}{\lambda\mu}\right)\right]=(\lambda\mu)^{p+n}T_{p,k+2}\left[\varphi(\mathbf{x})\right]+(\lambda\mu)^{p+n}(\ln\lambda+\ln\mu)T_{p,k+1}\left[\varphi(\mathbf{x})\right]$$

that the generalized function $T_{p,k}$ must be zero.

Statement 4A. Comparing expressions (3A) and (4A), we can conclude that the derivative of a generalized homogeneous function T_p of degree p with respect to homogeneity degree is a generalized first-order associated homogeneous function $T_{p,1}$ of degree p.

Indeed, differentiating equality (3A) with respect to parameter p (if such differentiation is possible) produces the following identity:

$$\frac{dT_p}{dp}\left[\varphi\left(\frac{\mathbf{x}}{\lambda}\right)\right] = \lambda^{p+n}\frac{dT_p}{dp}\left[\varphi(\mathbf{x})\right] + \frac{d\left(\lambda^{p+n}\right)}{dp}T_p\left[\varphi(\mathbf{x})\right] = \\ = \lambda^{p+n}\frac{dT_p}{dp}\left[\varphi(\mathbf{x})\right] + \lambda^{p+n}\ln\lambda\, T_p\left[\varphi(\mathbf{x})\right], \tag{6A}$$

which coincides with equality (4A) with an accuracy up to the notations.

Similarly, it is argued (without further substantiation of this statement) that the derivative of a generalized associated homogeneous function of order k with respect to homogeneity degree is always a generalized associated homogeneous function of order $k+1$.

Comment 3A. Unfortunately, this statement turns out to be incorrect for $k \geq 1$, if we take equality (5A) as the definition of a generalized associated function $T_{p,k}$ of order k. Differentiating condition (5A), we obtain a generalized function $T_{p,k+1}$ on the left-hand side of the equality, and a three-term expression involving generalized functions $T_{p,k+1}$, $T_{p,k}$ and $T_{p,k-1}$ on the right-hand side, since the coefficients λ^{p+n} and $\lambda^{p+n}\ln\lambda$ also depend on the parameter p. This three-term expression coincides with the two-term one (5A) for order $k+1$ if and only if the generalized function $T_{p,k-1}$ in equality (5P) is zero or if the order k is zero.

Similarly, given the validity of the assumption arising from this statement that a generalized associated homogeneous function $T_{p,k}$ of order k and degree p is the kth derivative of a homogeneous generalized function T_p with respect to the parameter p, after repeatedly differentiating the basic relation (3A) with respect to homogeneity degree p, more and more preceding generalized functions $T_{p,j}$ will appear in the corresponding functional equalities with nonzero coefficients as multipliers.

Statement 5A. In the future, constructing new generalized associated homogeneous functions of degree p and order k, the authors of [1] consistently apply differentiation with respect to the parameter p, as well as decomposition of the corresponding functionals into a Laurent series or a power series. In addition, in most cases the consideration is confined to the order $k=1$ for which the binomial identity (6P) that arises during differentiation holds true. In the general case, generalized functions obtained by differentiation are considered to be associated homogeneous functions of the respective order by the very fact that they were obtained, without additional verification of relation (5A).

Comment 4A. Thus, defining generalized associated homogeneous functions using relations (5A) turns out to be a dead-end in the reasoning both for the study in [1] and for the more detailed study in monograph [2]. In fact, this definition is not once applied in any way after it has been formulated. All calculations performed by the authors of [1, 2] turn out to be immaculately accurate, keeping intact the important and unique scientific results they have obtained, if differentiation with respect to the parameter p of generalized associated homogeneous functions of a lower order is used as a formal definition, and generalized associated homogeneous functions of zero order are considered generalized homogeneous functions ensuring that condition (3A) is fulfilled.

REFERENCES

1. **Gel'fand I. M., Shapiro Z. Y.,** Homogeneous functions and their extensions, Uspekhi Matem. Nauk. 10 (3(65)) (1955) 3–70 (in Russian).

2. **Gelfand I. M., Shilov G. E.,** Generailzed functions. Vol. 1. Properties and operations. American Mathematical Society (AMS) Chealsea Publishing, Providence (USA), 1964.

3. **Fikhtengol'ts G. M.,** The fundamentals of mathematical analysis, Vol. 1, Pergamon Press, Oxford, New York, 1965.

4. **Hobson E. W.,** The theory of spherical and ellipsoidal harmonics, Cambridge University Press, Cambridge, UK, 1931.

5. **Golikov Yu. K., Krasnova N. K.,** The generalized similarity principle and its application in electron spectrography, Appl. Phys. (2) (2007) 5–11 (in Russian).

6. **Golikov Y. K., Krasnova N. K.,** Application of electric fields uniform in the Euler sense in electron spectrography, Techn. Phys. 56 (2) (2011) 164–170.

7. **Golikov Yu. K., Krasnova N. K.,** Teoriya sinteza elektrostaticheskikh energoanalizatorov [The theory of synthesis of electrostatic energy analyzers], Polytechnic University Publishing, SPb, 2010 (in Russian).

8. **Golikov Yu. K., Krasnova N. K., Abramyonok O. A,** Electric spectrographs of charged particle flows with Euler-type potentials, Appl. Phys. (5) (2011) 69–73 (in Russian).

9. **Averin I. A., Berdnikov A. S., Gall N. R.,** The principle of similarity of trajectories for the motion of charged particles with different masses in electric and magnetic fields that are homogeneous in Euler terms, Tech. Phys. Lett. 43 (2) (2017) 156–158.

10. **Berdnikov A. S., Averin I. A., Krasnova N. K., Solovyev K. V.,** Theorems on the homogeneity of scalar and vector potentials for 3D electric and magnetic fields which are homogeneous in Euler terms, Usp. Prikl. Fiz. 5 (1) (2017) 10–27 (in Russian).

11. **Golikov Yu. K., Utkin K. G., Cheparukhin V. V.,** Raschet elementov elektrostaticheskikh elektronno-opticheskikh system [Calculation of elements of electrostatic electron-optical systems], Leningrad Polytechnic Institute Publishing, Leningrad, 1984 (in Russian).

12. **Gabdullin P. G., Golikov Yu. K., Krasnova N. K., Davydov S. N.,** The use of Donkin's formula in the theory of energy analyzers. I // Tech. Phys. 45 (2) (2000) 232–235.

13. **Gabdullin P. G., Golikov Yu. K., Krasnova N. K., Davydov S. N.,** Application of Donkin's formula in the theory of energy analyzers: Part II, Tech. Phys. 45 (3) (2000) 330–333.

14. **Averin I. A., Berdnikov A. S.,** Fringe fields of gridless electronic spectrographs with Euler's homogeneous electrostatic fields, Usp. Prikl. Fiz. 4 (1) (2016) 5–8. (In Russian)

15. **Berdnikov A. S., Averin I. A.,** A new approach to development of ion-optical systems for static mass spectrographs on the basis of the non-uniform Euler's homogeneous magnetic fields, Usp. Prikl. Fiz. 4 (1) (2016) 89–95 (in Russian).

16. **Berdnikov A. S., Averin I. A., Golikov Y. K.,** Static mass spectrometers of new type, using Euler's homogeneous electric and magnetic fields. I. General principle and single-stage systems, J. Anal. Chem. 71 (13) (2016) 1280–1287.

17. **Berdnikov A. S., Averin I. A., Golikov Y. K.,** Static mass spectrographs of a new type using Euler's homogeneous electric and magnetic fields. II: Conditions of high-order double focusing for two-cascade schemes, J. Anal. Chem. 71 (14) (2016) 1332–1340.

18. **Berdnikov A. S., Gall L. N., Antonov A. S., Solovyev K. V.,** Synthesis of fringing magnetic fields for static mass analyzers of the spectrographic type, J. Anal. Chem. 73 (14) (2018) 1301–1316.

19. **Golikov Yu. K., Berdnikov A. S., Antonov A. S., et al.,** Synthesis of electrode configurations that conserve fringing electric field homogeneity in Euler terms, Tech. Phys. 63 (4) (2018) 593–597.

20. **Golikov Yu. K., Berdnikov A. S., Antonov A. S., et al.,** Application of the Donkin formula in the theory of electrostatic prisms, Tech. Phys. 63 (11) (2018) 1659–1666.

21. **Golikov Yu. K., Berdnikov A. S., Antonov A. S., et al.,** Application of the Donkin formula in the theory of reflecting and turning devices, Tech. Phys. 64 (12) (2019) 1850–1865.

22. **Ivanov V. K.,** On the multiplication of homogeneous functions of several variables, Dokl. Akad. Nauk SSSR. 257 (1) (1981) 29–33 (in Russian).

23. **Ivanov V. K.,** Asymptotic approximation to a product of generalized functions, Soviet Math. (Izv. VUZ) 25 (1) (1981) 20–29.

24. **Estrada R., Kanwal R. P.,** The asymptotic expansion of certain series considered by Ramanujan

(Chapter), In book: Asymptotic analysis: A distributional approach, Birkhдuser Boston, Boston, USA (1994) 195–232.

25. **Danilov V. G., Maslov V. P., Shelkovich V. M.,** Algebras of the singularirties of singular solutions to first-order quasi-linear strictly hyperbolic systems, Theoret. and Math. Phys. 114 (1) (1998) 1–42.

26. **Estrada R., Kanwal R. P.,** A distributional approach to asymptotics: Theory and applications, 2-nd Ed., Birkhдuser Boston, Boston, USA, 2002.

27. **Albeverio S., Khrennikov A. Yu., Shelkovich V. M.,** Associated homogeneous *p*-adic distributions, Report 02133. MSI, Vaxjo University, Sweden, 2002.

28. **Albeverio S., Khrennikov A. Yu., Shelkovich V. M.,** Prisoyedinennyye odnorodnyye *p*-adicheskiye obobshchennyye funktsii [Associated homogeneous *p*-adic distributions], Dokl. RAN. 393 (3) (2003) 300–303 (in Russian).

29. **Albeverio S., Khrennikov A. Yu., Shelkovich V. M.,** Associated homogeneous *p*-adic distributions, J. Math. Anal. Appl. 313 (1) (2006) 64–83.

30**. Khrennikov A. Yu., Shelkovich V. M.,** Sovremennyy *p*-adicheskiy analiz i matematicheskaya fizika: teoriya i prilozheniya, [Modern *p*-adic analysis and mathematical physics: Theory and applications], Fizmatlit Publishing, Moscow, 2012 (in Russian).

31. **Shelkovich V. M.,** Associated and quasi associated homogeneous distributions (generalized functions), J. Math. Anal. Appl. 338 (1) (2008) 48–70.

32. **Albeverio S. A., Khrennikov A. Yu., Shelkovich V. M.,** Non-linear singular problems of *p*-adic analysis: Associative algebras of *p*-adic distributions, Izvestiya: Mathematics. 69 (2) (2005) 221–263.

33. **Albeverio S., Khrennikov A. Yu., Shelkovich V. M.,** Theory of *p*-adic distributions: Linear and nonlinear models (London Mathematical Society Lecture Note Series. Vol. 370), Cambridge University Press, Cambridge, 2010.

34. **Kreyn S. G.,** Functional analysis. (Reference book), Library of University of Michigan, 1967 (Chap. X, §2, P. 8).

35. **Von Grudzinski O.,** Quasihomogeneous distributions (Book Series: Mathematics Studies, Vol. 165), Elsevier, Amsterdam (North-Holland), 1991.

36. **Kuczma M.,** An introduction to the theory of functional equations and inequalities: Cauchy's equation and Jensen's inequality, Birkhдuser Basel, Basel, Switzerland, 2009.

37. **Berdnikov A. S., Solovyev K. V., Krasnova N. K.,** Mutually homogeneous functions with finite-sized matrices, St. Petersburg Polytechnical State University Journal. Physics and Mathematics. 13 (1) (2020) 42–53.

38. **Berdnikov A. S., Solovyev K. V., Krasnova N. K.,** Chains of fundamental mutually homogeneous functions with a common real eigenvalue, St. Petersburg Polytechnical State University Journal. Physics and Mathematics. 13 (2) (2020) 53–71 (in Russian).

39. **Berdnikov A. S., Solovyev K. V., Krasnova N. K.,** General formulas for chains of fundamental mutually homogeneous functions with a common pair of complex conjugate eigenvalues, St. Petersburg Polytechnical State University Journal. Physics and Mathematics. 13 (2) (2020) 72–88 (in Russian).

СПИСОК ЛИТЕРАТУРЫ

1. **Гельфанд И. М., Шапиро З. Я.** Однородные функции и их приложения // Успехи математических наук. 1955. Т. 65) 3 № .10). С. 70–3.

2. **Гельфанд И. М., Шилов Г. Е.** Обобщенные функции и действия над ними (Серия «Обобщенные функции». Вып. 1). М.: Гос. изд-во физ.-мат. лит-ры, 1959. 472 с.

3. **Фихтенгольц Г. М.** Курс дифференциального и интегрального исчисления. В 3 тт. -8е изд. Т. 1. М.: Физматлит, 607 .2003 с.

4. **Гобсон Е. В.** Теория сферических и эллипсоидальных функций. Пер. с англ. М.: Изд-во иностранной литературы, 476 .1952 с.

5. **Голиков Ю. К., Краснова Н. К.** Обобщенный принцип подобия и его применение в электронной спектрографии // Прикладная физика. 2 № .2007. С. 11–5.

6. **Голиков Ю. К., Краснова Н. К.** Электрические поля, однородные по Эйлеру, для электронной спектрографии // Журнал технической физики. 2011. Т. 2 № .81. С. 15–9.

7. **Голиков Ю. К., Краснова Н. К.** Теория синтеза электростатических энергоанализаторов. СПб.: Изд-во Политехнического университета, 409 .2010 с.

8. **Голиков Ю. К., Краснова Н. К., Абрамёнок О. А.** Электрические спектрографы потоков заряженных частиц с потенциалами эйлерова типа // Прикладная физика. 5 № .2011. С. 73–69.

9. **Аверин И. А., Бердников А. С., Галль Н. Р.** Принцип подобия траекторий при движении заряженных частиц с разными массами в однородных по Эйлеру электрических и магнитных полях // Письма в Журнал технической физики. 2017. Т. 3 № .43. С. 43–39.

10. **Бердников А. С., Аверин И. А., Краснова Н. К., Соловьев К. В.** Об однородности скалярных и векторных потенциалов электрических и магнитных полей, однородных по Эйлеру // Успехи прикладной физики. 2017. Т. 1 № .5. С. 27–10.

11. **Голиков Ю. К., Уткин К. Г., Чепарухин В. В.** Расчет элементов электростатических электронно-оптических систем. Ленинград: Изд-во Ленинградского политехнического института, 79 .1984 с.

12. **Габдуллин П. Г., Голиков Ю. К., Краснова Н. К., Давыдов С. Н.** Применение формулы Донкина в теории энергоанализаторов. I // Журнал технической физики. 2000. Т. 70. № 2. С. 91–94.

13. **Габдуллин П. Г., Голиков Ю. К., Краснова Н. К., Давыдов С. Н.** Применение формулы Донкина в теории энергоанализаторов. II // Журнал технической физики. 2000. Т. 70. № 3. С. 44–47.

14. **Аверин И. А., Бердников А. С.** Краевые поля бессеточных электронных спектрографов с однородными по Эйлеру электростатическими полями // Успехи прикладной физики. 2016. Т. 1 № .4. С. 8–5.

15. **Бердников А. С., Аверин И. А.** Новый подход к разработке ионно-оптических схем статических масс-спектрографов на основе неоднородных магнитных полей, однородных по Эйлеру // Успехи прикладной физики. 2016. Т. 1 № .4. С. 95–89.

16. **Бердников А. С., Аверин И. А., Голиков Ю. К.** Статические масс-спектрографы нового типа, использующие электрические и магнитные поля, однородные по Эйлеру. I. Общий принцип и однокаскадные схемы // Масс-спектрометрия. 2015. Т. 12. № 4. С. 272–281.

17. **Бердников А. С., Аверин И. А., Голиков Ю. К.** Статические масс-спектрографы нового типа, использующие электрические и магнитные поля, однородные по Эйлеру. II. Условия двойной фокусировки высокого порядка у двухкаскадной схемы // Масс-спектрометрия. 2016. Т. 13. № 1. С. 11–20.

18. **Бердников А. С., Галль Л. Н., Антонов А. С., Соловьев К. В.** Синтез краевых магнитных полей для статических масс-анализаторов спектрографического типа // Масс-спектрометрия. 2018. Т. 1 № .15. С. 43–26.

19. **Голиков Ю. К., Бердников А. С., Антонов А. С., Краснова Н. К., Соловьев К. В.** Синтез электродных конфигураций, сохраняющих для краевых электрических полей свойство однородности по Эйлеру // Журнал технической физики. 2018. Т. 4 № .88. С. 613–609.

20. **Голиков Ю. К., Бердников А. С., Антонов А. С., Краснова Н. К., Соловьев К. В.** Применение формулы Донкина в теории электростатических призм // Журнал технической физики. 2018. Т. 11 № .88. С. 1719–1711.

21. **Голиков Ю. К., Бердников А. С., Антонов А. С., Краснова Н. К., Соловьев К. В.** Применение формулы Донкина в теории отражающих и поворотных устройств // Журнал технической физики. 2019. Т. 12 № .89. С. 1964–1947.

22. **Иванов В. К.** Об умножении однородных функций нескольких переменных // Доклады Академии наук СССР. 1981. Т. 1 № .257. С. 33–29.

23. **Иванов В. К.** Асимптотическое приближение к произведению обобщенных функций // Известия высших учебных заведений. Математика. 1981. Т. 25. № 1. С. 20–29.

24. **Estrada R., Kanwal R. P.** The asymptotic expansion of certain series considered by Ramanujan (Chapter) // Asymptotic analysis: A distributional approach. Boston, USA: Birkhduser Boston, 1994. Pp. 195–232.

25. **Данилов В. Г., Маслов В. П., Шелкович В. М.** Алгебры особенностей сингулярных решений квазилинейных строго гиперболических систем первого порядка // Теоретическая и математическая физика. 1998. Т. 114. № 1. С. 3–55.

26. **Estrada R., Kanwal R. P.** A distributional approach to asymptotics: Theory and applications. Second Edition. Boston, USA: Birkhduser Boston, 2002. 454 p.

27. **Albeverio S., Khrennikov A. Yu., Shelkovich V. M.** Associated homogeneous *p*-adic distributions. Report 02133. MSI. Sweden, Vaxjo University, 2002.

28. **Альбеверио С., Хренников А. Ю., Шелкович В. М.** Присоединенные однородные *p*-адические обобщенные функции // Доклады Российской академии наук. 2003. Т. 393. № 3. С. 300–303.

29. **Albeverio S., Khrennikov A. Yu., Shelkovich V. M.** Associated homogeneous *p*-adic distributions // Journal of Mathematical Analysis and Applications. 2006. Vol. 313. No. 1. Pp. 64–83.

30. **Хренников А. Ю., Шелкович В. М.** Современный p-адический анализ и математическая физика: теория и приложения. М.: Физматлит, 2012. 452 с.

31. **Shelkovich V. M.** Associated and quasi associated homogeneous distributions (generalized functions) // Journal of Mathematical Analysis and Applications. 2008. Vol. 338. No. 1. Pp. 48–70.

32. **Альбеверио С. А., Хренников А. Ю., Шелкович В. М.** Нелинейные сингулярные проблемы *p*-адического анализа: ассоциативные алгебры *p*-адических распределений // Известия Российской академии наук. Серия Математическая. 2005. Т. 69. № 2. С. 3–44.

33. **Albeverio S., Khrennikov A. Yu., Shelkovich V. M.** Theory of *p*-adic distributions: Linear and nonlinear models (London Mathematical Society Lecture Note Series. Vol. 370). Cambridge: Cambridge University. Press, 2010. 351 p.

34. Функциональный анализ. Справочник. 2-е изд. Ред. С. Г. Крейн (Справочная математическая библиотека). М.: Наука, 1972. 544 с. С. 485.

35. **Von Grudzinski O.** Quasihomogeneous distributions (Mathematics Studies. Vol. 165). Amsterdam (North-Holland): Elsevier, 1991. 448 p.

36. **Kuczma M.** An introduction to the theory of functional equations and inequalities**:** Cauchy's equation and Jensen's inequality, Basel, Switzerland: Birkhduser Basel, 2009. 595 p.

37. **Бердников А. С., Соловьев К. В., Краснова Н. К.** Взаимно-однородные функции с матрицами конечного размера // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2020. Т. 1 № .13. С. 53–42.

38. **Бердников А. С., Соловьев К. В., Краснова Н. К.** Цепочки фундаментальных присоединенных однородных функций с общим вещественным собственным числом // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2020. Т. 2 № .13. С. 71–53.

39. **Бердников А. С., Соловьев К. В., Краснова Н. К.** Общие формулы для цепочек фундаментальных присоединенных однородных функций с общей парой комплексно-сопряженных собственных чисел // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2020. Т. 13. № 2. С. 72–88.

THE AUTHORS

BERDNIKOV Alexander S.
Institute for Analytical Instrumentation of the Russian Academy of Sciences
31–33 Ivan Chernykh St., St. Petersburg, 198095, Russia
asberd@yandex.ru
ORCID: 0000-0003-0985-5964

BULYANITSA Anton L.
Institute for Analytical Instrumentation, RAS
Peter the Great St. Petersburg Polytechnic University
31–33 Ivan Chernykh St., St. Petersburg, 198095, Russia
antbulyan@yandex.ru
ORCID: 0000-0002-9235-8549

SOLOVYEV Konstantin V.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
k-solovyev@mail.ru
ORCID: 0000-0003-3514-8577

СВЕДЕНИЯ ОБ АВТОРАХ

БЕРДНИКОВ Александр Сергеевич – *доктор физико-математических наук, заведующий лабораторией, главный научный сотрудник Института аналитического приборостроения Российской академии наук.*
198095, Россия, г. Санкт-Петербург, ул. Ивана Черных, 31–33, лит. А.
asberd@yandex.ru
ORCID: 0000-0003-0985-5964

БУЛЯНИЦА Антон Леонидович – *доктор физико-математических наук, ведущий научный сотрудник Института аналитического приборостроения Российской академии наук, профессор кафедры высшей математики Санкт-Петербургского политехнического университета Петра Великого.*
198095, Россия, г. Санкт-Петербург, ул. Ивана Черных, 31–33, лит. А.
antbulyan@yandex.ru
ORCID: 0000-0002-9235-8549

СОЛОВЬЕВ Константин Вячеславович – *кандидат физико-математических наук, доцент Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого, младший научный сотрудник Института аналитического приборостроения Российской академии наук.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
k-solovyev@mail.ru
ORCID: 0000-0003-3514-8577

Received 22.09.2023. Approved after reviewing 16.02.2024. Accepted 16.02.2024.
Статья поступила в редакцию 22.09.2023. Одобрена после рецензирования 16.02.2024. Принята 16.02.2024.

EXPERIMENTAL TECHNIQUE AND DEVICES

Original article
DOI: https://doi.org/10.18721/JPM.17205

A MICROWAVE METHOD FOR MEASURING THE LOW-FREQUENCY NOISE OF TRANSISTORS

V. G. Usychenko[1], *A. S. Chernova*[2,1]✉

[1] JSC "Svetlana-Electronpribor", St. Petersburg, Russia;

[2] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ chernova-nastya@mail.ru

Abstract. In the article, we have proposed to use microwave-fluctuation meters resistant to external intense electromagnetic noise in order to measure the low-frequency (LF) noise of microwave transistors working under these conditions. The transistor located on the board is excited by a low-noise microwave generator, the oscillation amplitude of which, being modulated by the LF noise of the transistor, is measured by a microwave spectrum analyzer. The proposed method was tested on GaN/AlGaN heterotransistors, in whose channels the electron density was formed by spontaneous and piezoelectric polarization. In addition to experimental testing, a theoretical justification for the method is presented. We obtained conditions in which the normalized spectra of oscillation amplitude fluctuations were similar to the normalized LF noise of the transistor current.

Keywords: low frequency noise, microwave transistor, fluctuations of microwave oscillations, GaN/AlGaN heterostructure

Citation: Usychenko V. G., Chernova A. S., A microwave method for measuring the low-frequency noise of transistors, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 61–79. DOI: https://doi.org/10.18721/JPM.17205

Научная статья
УДК 621.391.822, 621.391.823
DOI: https://doi.org/10.18721/JPM.17205

МИКРОВОЛНОВЫЙ МЕТОД ИЗМЕРЕНИЯ НИЗКОЧАСТОТНЫХ ШУМОВ ТРАНЗИСТОРОВ

В. Г. Усыченко[1], *А. С. Чернова*[2,1]✉

[1] АО «Светлана – Электронприбор», Санкт-Петербург, Россия;

[2] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉ chernova-nastya@mail.ru

Аннотация. Для измерения низкочастотных (НЧ) шумов СВЧ-транзисторов, работающих в условиях внешних интенсивных электромагнитных помех, предлагается использовать устойчивые к их воздействию измерители флуктуаций СВЧ-колебаний. Транзистор, находящийся на плате, возбуждают малошумящим СВЧ-генератором, амплитуда колебаний которого, модулированная НЧ-шумами транзистора, измеряется СВЧ-анализатором спектра. Методика опробована на гетеротранзисторах GaN/AlGaN, в каналах которых плотность электронов формировалась посредством спонтанной и пьезоэлектрической поляризации без какого-либо дополнительного легирования. Помимо экспериментального тестирования, представлено теоретическое обоснование предложенного метода. Получены условия, при которых нормированные спектры флуктуаций амплитуды колебания аналогичны нормированным НЧ-шумам тока транзистора.

Ключевые слова: низкочастотный шум, СВЧ-транзистор, низкочастотные помехи, флуктуации СВЧ-колебаний, гетероструктура GaN/AlGaN

Ссылка для цитирования: Усыченко В. Г., Чернова А. С. Микроволновый метод измерения низкочастотных шумов транзисторов // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 2. С. 61–79. DOI: https://doi.org/10.18721/JPM.17205

Introduction

Since the 1950s, low-frequency (LF) noise whose spectrum is localized in the $0 < F < 10^6$ Hz frequency range has served to provide insights into structural defects in semiconductor crystals and various types of devices [1–7]. In practice, LF noise limits the sensitivity of sensors as well as other measuring and recording instruments operating in a wide variety of frequency ranges. For example, the detection accuracy of Doppler radars for low-slow-small (LSS) targets is largely limited by low-frequency noise of transistors used in microwave generators, mixers and amplifiers. New semiconductor materials and devices based on them appear as constant advances are made in semiconductor technology. There is also continued interest towards low-frequency noise and its origin; measuring the level of this noise remains a crucial problem.

However, measurement and analysis of this noise, especially in operating facilities, are greatly hindered by numerous intense electromagnetic interference generated externally in the LF range by various technical and industrial devices. The experience accumulated indicates that the techniques that previously proved successful for shielding and reduction of interference [8] can no longer combat the problem, so it is increasingly difficult to use LF instrumentation without additional measures. Practical evidence suggests that highly sensitive microwave instrumentation is better protected from LF interference. Due to operational need, we previously tested whether such devices were capable of measuring LF noise in microwave transistors. These experiments were successful.

The goal of this study is to practically test the technique for measuring LF noise in microwave transistors and provide a theoretical justification for this technique.

Problem statement and solution

One particular case when it is necessary to measure LF noise in transistors is the design of ultra-low-noise microwave devices such as amplifiers. This type of noise is measured when the transistor operates in DC mode, switching from one set of key parameters (drain voltage U_d, drain current I_d, gate voltage U_g) to another. The measurement results are then summarized to predict the inevitable fluctuations in the amplitude and phase of microwave oscillations of a self-excited generator or amplifier incorporated into such a transistor.

We propose the reverse approach, i.e., we consider a prototype microwave amplifier installed on the given transistor, feed a low-noise microwave oscillation to its input, and then measure the energy spectrum of amplitude fluctuations for this oscillation, generated by transistor noise. In many cases, this sequence of operations allows to extract a sufficient amount of data for practical purposes. However, it is also useful to obtain additional information, establishing the relationship of amplitude fluctuations with the mean (direct) current noise

This allows to conduct dynamic rather than static measurements of noise, where the transistor parameters U_d, U_g, I_d vary nonlinearly along with microwave oscillations. We simultaneously obtain information about LF noise of the transistor (averaged over the oscillation period) both in microwave and DC mode.

The practical implementation of this procedure was as follows: a transistor mounted on a microwave PCB (printed circuit board) was connected to low-noise gate-bias and drain-bias sources. A bandpass filter was located on the board in the drain circuit, tuned to a carrier frequency of 3 GHz $= f = 2\pi/\omega >> F$; here F is the high-frequency cutoff, exceeding the carrier frequency f by orders of magnitude.

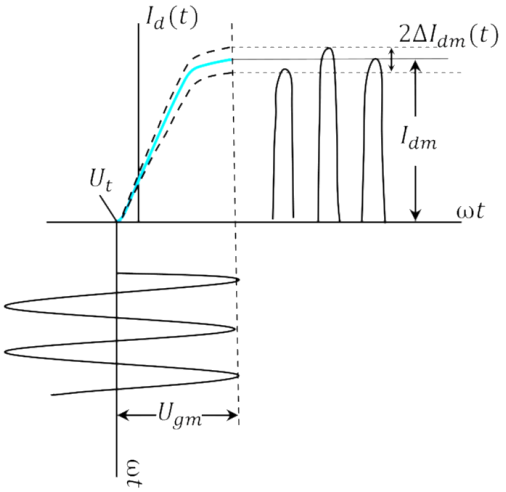

Fig. 1. Generation of amplitude-modulated drain current pulse train in nonlinear mode: time dependences of voltage U_t and I_d, transfer characteristic (cyan curve) of the transistor at constant U_d as well as other key parameters of the process are shown

The simplified transfer characteristic of the transistor $I_d(U_g)\big|_{U_d=const}$, measured at constant voltage U_d, is shown in Fig. 1 by a solid line. The LF noise sources in the transistor modulate the drain current, so the transfer characteristic fluctuates deviating from its mean level (this is schematically shown by the dashed line in Fig. 1). A DC bias voltage $U_g = U_t$, equal to the threshold value of U_t, is applied to the gate. A highly stable (carrier) microwave oscillation $U_{gm}\cos\omega t$ is applied to the gate along with the DC bias voltage; its amplitude $U_{gm} < U_d$ is sufficiently large, and the transistor operates in saturation. This nonlinear mode is typical for self-excited generators and microwave power amplifiers. The shape of the drain current differs significantly from the sine wave applied to the gate. Instead, a periodic train of unipolar pulses is generated (see Fig. 1), similar to a trapezoid or rectangular pulse with a rounded peak.

This sequence can be expanded into a trigonometric Fourier series:

$$I_d(t) = I_{dm}\left(\frac{c_0}{2} + c_1\cos\omega t + c_2\cos\omega t + c_3\cos\omega t + ...\right) = \\ = I_{dm0} + I_{dm1}\cos\omega t + I_{dm2}\cos 2\omega t + I_{dm3}\cos 3\omega t + ..., \quad (1)$$

where I_{dm} is the amplitude of the pulse, I_{dm0} is the DC component of the current; I_{dm1}, I_{dmj} are the amplitudes of the carrier frequency and the jth harmonics of the current, respectively; c_j are the series coefficients.

LF fluctuations of the transfer characteristic in turn induce fluctuations in the current pulse amplitude, expressed as

$$I_{dm}(t) = \overline{I}_{dm} + \Delta I_{dm}(t) = \overline{I}_{dm}(1+\alpha(t)), \quad (2)$$

where $\overline{I}_{dm}$ is the average current amplitude; $\Delta I_{dm}(t)$ are the low-frequency amplitude fluctuations; $\alpha(t)$ are the relative amplitude fluctuations, $\alpha(t) = \Delta I_{dm}(t)/\overline{I}_{dm}$.

Since the frequency of LF noise F is several orders of magnitude lower than the carrier frequency f ($F << f$), fluctuations are very slow compared to the oscillation period.

Substituting expression (2) into Eq. (1), we obtain the following series:

$$I_d(t) = (1+\alpha(t))\left[\overline{I}_{dm0} + \overline{I}_{dm1}\cos\omega t + \overline{I}_{dm2}\cos 2\omega t + \overline{I}_{dm3}\cos 3\omega t + ...\right]. \quad (3)$$

As the bandpass filter located in the drain circuit eliminates all microwave harmonics except the first-order and the zeroth-order one, only the fluctuations of the DC component of the current and the carrier remain. As a result, we obtain the following expression:

$$I_d(t) = (1+\alpha(t))\left[\overline{I}_{dm0} + \overline{I}_{dm1}\cos\omega t\right]. \quad (4)$$

Both currents on the right-hand side of the equality fluctuate by the same law. Let us measure the carrier current fluctuations.

To avoid complicating the formulas, let us consider the case when, instead of noise $\alpha(t)$ with continuous spectrum, only one of its spectral components, $\alpha_{\Delta F}(t)$, located at frequency F in a

narrow band $\Delta F << F$, influences the oscillations. The narrowband noise component can be replaced by an oscillation with a slowly varying amplitude $b(t)$ and phase [9], i.e.,

$$\theta(t) : \alpha_{\Delta F}(t) = b(t)\cos\left[\Omega t + \theta(t)\right],$$

where $\Omega = 2\pi F$.

If we ignore the phase fluctuations, which do not play a significant role in the considered example, we obtain:

$$I_{dm}(t) = \overline{I}_{dm0}\left[1 + b(t)\cos\Omega t\right] + \overline{I}_{dm1}\left[1 + b(t)\cos\Omega t\right]\cos\omega t =$$

$$= \overline{I}_{dm0}\left[1 + b(t)\cos\Omega t\right] + \overline{I}_{dm1}\left[\cos\omega t + \frac{b(t)}{2}\cos(\omega + \Omega)t + \frac{b(t)}{2}\cos(\omega - \Omega)t\right].$$

Noise is typically measured in units of power. The amplitude of narrowband DC fluctuations is equal to

$$\Delta I_{dm0}(t) = b(t)\overline{I}_{dm0},$$

which means that its power (with an accuracy up to the resistance) follows the expression

$$P_0(F) = b^2\overline{I}^2_{dm0} / 2;$$

and the electric power $P_0 = \overline{I}^2_{dm0}$.

The ratio of these powers is expressed by the formula

$$\frac{P_0(F)}{P_0} = \frac{\Delta I^2_{dm0}}{\overline{I}^2_{dm0}} = b^2 / 2. \tag{5}$$

The microwave oscillation consists of a carrier oscillation and two modulation sidebands with the amplitudes $b\overline{I}_{dm1} / 2$. The carrier power is equal to $P_1 = \overline{I}^2_{dm1} / 2$, and the power of the two sidebands is

$$P_s(F) = \frac{\overline{I}^2_{dm1}b^2}{8} + \frac{\overline{I}^2_{dm1}b^2}{8} = \frac{\overline{I}^2_{dm1}b^2}{4}.$$

The power ratio

$$\frac{P_s(F)}{P_{1\ 1}} = \frac{\Delta I^2_{dm1}}{\overline{I}^2_{dm1}} = b^2 / 2$$

turns out to be the same as for the currents (see Eq. (5)):

$$\frac{\Delta I^2_{dm0}}{\overline{I}^2_{dm0}} = \frac{\Delta I^2_{dm1}}{\overline{I}^2_{dm1}}.$$

Thus, the power of the relative fluctuations in the amplitude of the two sidebands of the carrier frequency.

Moving from narrowband fluctuation to continuous spectrum, we obtain:

$$\frac{S^2_{i0}(F)}{\overline{I}^2_{dm0}} = \frac{S^2_{i1}(F)}{\overline{I}^2_{dm1}} = \delta S^2_i(F), \tag{6}$$

where $S_{i0}^2(F)$, $S_{i1}^2(F)$ are the power spectral densities (PSDs) of LF noise from direct current and sidebands of the carrier frequency, respectively.

We replace these two equal spectral densities by a single quantity $\delta S_i^2(F)$; this is the spectral density of relative intensity noise (RIN), normalized by the rms current, measured in Hz^{-1}. Non-normalized noise spectral density $S_{i0}^2(F)$ is also often considered in practice, measured in A^2/Hz.

Notably, the normalized PSDs of amplitude noise in amplifiers are generally measured in decibels relative to the power of the carrier frequency in 1-Hz bandwidth, where the measurement unit is dBn/Hz. For example, the relative amplitude fluctuations measured with a carrier frequency offset F equal to

$$\delta S_i^2(F) = 2 \cdot 10^{-14} \text{ Hz}^{-1},$$

which, in measurement units adopted in microwave technology, equals

$$10 \lg(2 \cdot 10^{-14}) = 3 - 140 = -137 \text{ dBc/Hz}.$$

Turning to the measurement procedure, we should note that instead of the microwave current of the amplifier appearing in the equations above, it is the output microwave voltage with fluctuating amplitude that is applied to the input of the microwave spectrum analyzer:

$$U(t) = \left[1 + \upsilon(t)\right] \bar{U} \cos \omega t,$$

where $\bar{U} = \bar{I}_{dm1} \cdot R_e$ is the mean amplitude of the microwave voltage, and

$$\upsilon(t) = \frac{\Delta I_{dm1}(t)}{\bar{I}_{dm1}} R_e \tag{7}$$

are the relative fluctuations in the amplitude of the microwave voltage of the amplifier.

R_e in these expressions is the equivalent resistance of the bandpass filter located in the drain circuit, $R_e = \rho Q$; it is typically equal to the wave resistance of the microwave circuit on the board, R_e = 50 ohms. Such equalities $R_e = \rho Q$ = 50 ohms are obtained by selecting the Q factor, the wave resistance ρ of the filter and the coupling parameters at the carrier frequency.

Agreement between theory and practice should be observed in noise measurements. In theory, the amplitude of the carrier current exceeds the amplitudes of other harmonics; this condition must be also satisfied in practice. The shape of the current pulse influencing the distribution of the amplitudes of Fourier harmonics does not fundamentally change if the following inequality is satisfied:

$$U_d - U_{gm} > U_{dm1} = I_{dm1} R_e. \tag{8}$$

Otherwise, rather than flow from the source to the drain at times when the oscillation amplitudes at the gate and drain are in opposite phase, the current flows in a more complex pattern. This can significantly change the ratio between the amplitudes of harmonics, cause misalignment of circuits and distort the measurement results.

Measurements of HFET characteristics and results

The technique described above was used to study LF noise in GaN/AlGaN-based heterostructure field-effect transistors (HFETs), with the density of two-dimensional electron gas in the channel formed by spontaneous and piezoelectric polarization [10] without any additional doping. These are high-power transistors assembled on a silicon carbide substrate by molecular beam epitaxy (MBE) technology. They operate at frequencies $f \le 6$ GHz, have a gate length of 0.5 μm, and a source-drain gap of 4.8 μm. According to the manufacturer's specifications, the transistor operating at a frequency f = 3 GHz amplifies the power P_{in} = 100 MW of the input microwave signal by 25 times with an efficiency coefficient Eff = 65% at a drain voltage U_d = 28 V. The measurement circuit for the transistor noise is shown in Fig. 2. Three transistor models were examined, yielding similar measurement results.

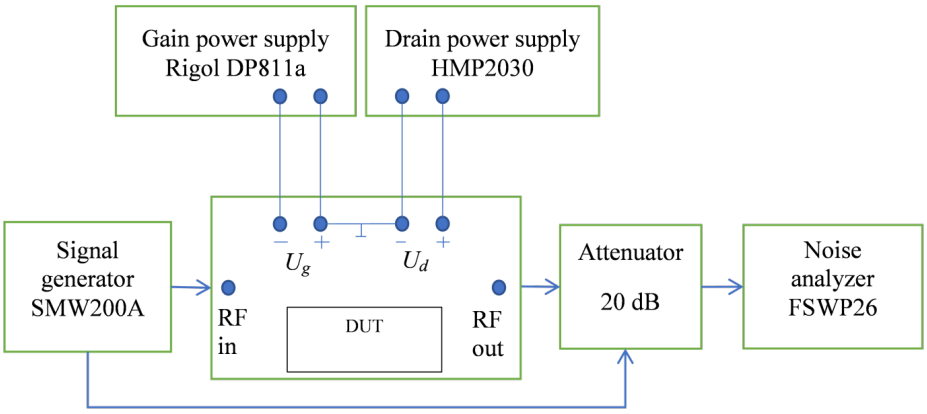

Fig. 2. Measurement circuit

The signal from a low-noise SMW200A generator is applied to the microwave input of the transistor considered; the oscillation amplified by the transistor is applied to an R&S FSWP26 noise analyzer equipped with an attenuator to prevent overload. Rigol DP811a and R&S HMP2030 power supplies were used to power the drain and gate of the transistor, each supplemented with filters suppressing their intrinsic LF noise and electrical interference. An HMP2030 display was used to monitor voltages and currents in the gate and drain circuits. Voltage monitoring in the drain circuit was carried out with an external Agilent U3401A voltmeter, which was disconnected from the circuit during noise measurements.

The power of the SMW200A low-noise microwave generator used did not exceed 63 MW, which is equivalent to an amplitude $U_{gm} = 2.5$ V at a load resistance $R = 50$ ohms. The transistor's transconductance, which was initially very high, reaching 0.5 A/V and rapidly transitioning to saturation, provided the drain current with a shape similar to that shown in Fig. 1 at such an amplitude.

The gate bias U_g was equal to the threshold value U_T. Since U_T changes when the drain voltage is adjusted, it was determined by varying the gate voltage U_g so that the drain current I_{dT} at the nominal value $U_d = 28$ V was about three orders of magnitude less than the highest permissible current of the transistor. The value $U_g = U_T = -3.14$ V was obtained at $I_{dT} = 1$ mA. The value $U_g - U_T = 0$ was maintained constant during noise measurements at any level of U_d, which was limited only by the dissipated power of the transistor $P = 1.5$ W at high excitation amplitudes of U_{gm} (this is shown in the legends to the figure).

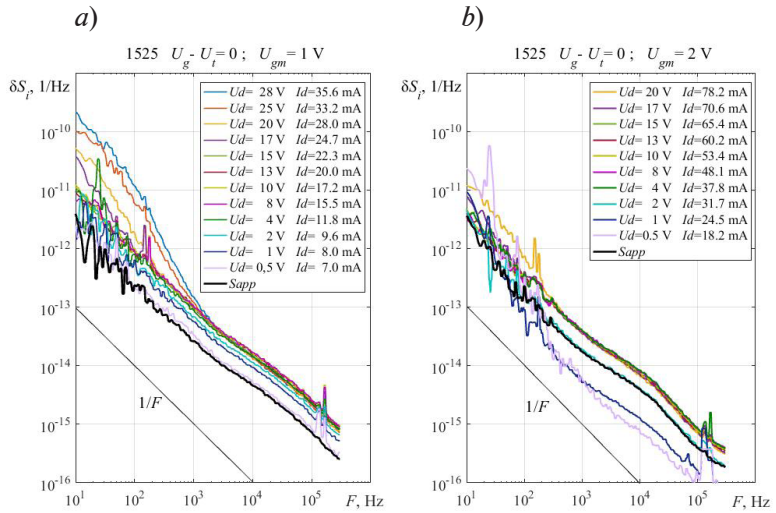

Fig. 3. Measured frequency dependences for relative fluctuations in transistor drain current at 1 V (*a*) and 2 V (*b*) amplitudes of U_{gm} excitation.
Black curve corresponds to the instrumentation's intrinsic noise PSD (Sapp)

Fig. 3 shows the experimental dependences for drain-current RIN spectral density in transistor 1525 excited by microwave oscillations with amplitudes U_{gm} equal to 1 and 2 V. The sharpened peaks on the spectra are caused by parasitic interference. The bold black curve corresponds to the instrumentation's intrinsic noise PSD. These noises were determined in accordance with the manual by analyzer (see Fig. 2), bypassing the transistor. However, the input oscillation noise is reduced in nonlinear mode of the transistor due to saturation of the transfer characteristic (see Fig. 1). We did not take this phenomenon into account, so the instrumentation noise was somewhat overestimated in our experiments.

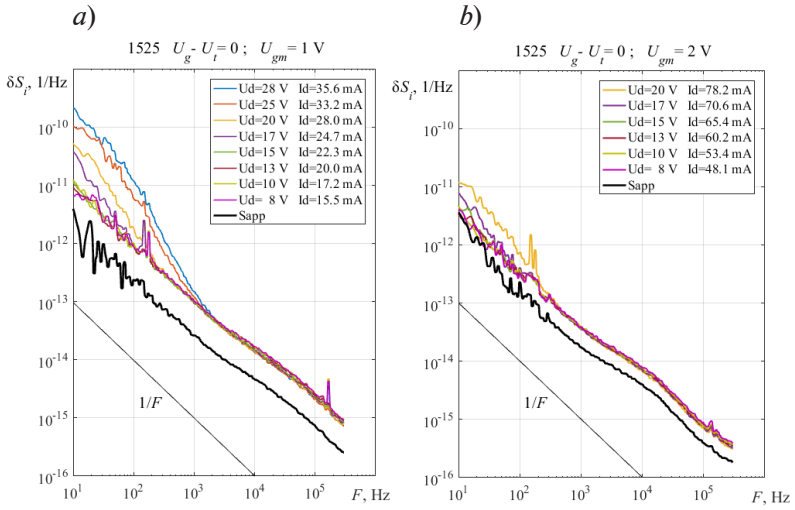

Fig. 4. Dependences similar to those in Fig. 3, with distortions removed

It can be seen from Fig. 3 that part of the measured spectra lies below the noise level of the instrumentation. Such distortions appear because condition (8) is not satisfied. The dependences with the distortions removed are shown in Fig. 4. Analyzing the data, we can conclude that a twofold increase in the amplitude U_{gm} of the exciting oscillation at constant U_d leads to an approximately three-fold increase in the drain current and, accordingly, to increased power dissipated by the transistor. To prevent transistor overloading in U_{gm} = 2 V mode, the maximum drain voltage was limited to U_d = 20 V.

Results and discussion

Let us analyze the experimental data presented in the previous section. The normalized PSDs of transistor current noise measured in the voltage ranges 8 V $\le U_d \le$ 15 V and current ranges 15.5 mA $\le I_d \le$ 65.4 mA (Fig. 4) overlap, varying as $1/F$ with frequency tuning (a small rise in the high-frequency region is caused by instrumentation noise added to the curves from below). The noise PSDs at a frequency F = 1 kHz, measured at the amplitudes of exciting vibrations U_{gm} = 1 and 2 V and calculated taking into account the instrumentation noise (see Fig. 4, *a* and *b*), amounted to $77.5 \cdot 10^{-14}$ and $10 \cdot 2.0^{-14}$ Hz^{-1}, respectively. Thus, the noise level in the second case is almost four times lower than the level obtained in the first case. The overlapping of normalized noise PSDs measured at different currents implies that they are generated by the same source of noise.

The results obtained are consistent with the data published by other authors (see [11–13]) for microwave high-electron-mobility transistors (HEMTs).

However, at elevated drain voltages, when $U_d >$ 15 V, the noise begins to increase at the analyzed frequencies $F <$ 1 kHz, reaching progressively higher levels the higher the value of U_d. The shape of its elevated spectra is characteristic for generation-recombination noise [1–4], gradually tending back towards the $1/F$-type dependence with decreasing frequency, but at a higher level. A source of LF noise of a different physical nature prevails in this frequency range.

Let us first discuss the possible nature of primary ($1/F$) noise, starting with the simpler case when microwave oscillation is not applied to the gate.

The drain current does not appear at a bias voltage $U_g - U_T \le 0$, the height of the gate gap through which electrons drift from the source to the drain is zero, which means that the resistance R of the gate region of the channel can be considered infinitely large. As the voltage increases, the gap opens slightly at $U_g - U_T > 0$, its height increases, the resistance R decreases, and the current flowing through the channel increases. If the resistance

$$R(t) = \overline{R} + \Delta R(t) = \overline{R}(1 + r(t))$$

fluctuates, its fluctuations $\Delta R(t)$, generated by the processes also occurring in the transistor in the absence of current, modulate the drain current. These fluctuations can be generated by random interactions of electrons in the gate region of the channel with donor traps [13], whose energy levels are below the Fermi level, mainly in the GaN layer under the channel. Relative resistance fluctuations $r(t) = \Delta R(t) / \overline{R}$ do not depend on current, so the RIN spectral densities $\delta S_i^2(F)$ also do not depend on current: they overlap each other (see Fig. 4).

In our case, when microwave oscillation with an amplitude U_{gm} is superimposed on the DC bias voltage $U_g - U_T = 0$, the height of the gate gap corresponds to its height averaged over the oscillation period. Then the higher the U_{gm} value, the greater the gap height and the lower the $1/F$ noise level. This is exactly the behavior exhibited by $1/F$-noise shown in Fig. 4: the noise PSD in Fig. 4,*a* is almost four times lower than in Fig. 4,*b*. Noise reduction with an increase in potential difference $U_g - U_T > 0$ is typical for HEMTs [12, 13].

Now consider the source of generation-recombination noise (GRN) appearing at frequencies $F \le 1$ kHz at elevated drain voltages $U_d > 15$ V. If the GRN source is localized in any small region of a semiconductor device, then its RIN spectral density is proportional to the rms current flowing through this region. At the same time, the PSD normalized by the rms current does not depend on the current. A localized source of $1/F$ noise behaves similarly: as evident from Fig. 4, its spectral densities, measured at different currents, overlap each other. However, the GRN under consideration behaves differently. It can be seen from Fig. 4,*a* that GRN starts to appear at currents $I_d > 15.3$ mA and their RIN increases with increasing U_d.

Understanding the physical nature of this noise source is a separate complex problem beyond the scope of this study.

Conclusion

The considered microwave technique allows to measure the low-frequency noise of transistors under conditions of intense electromagnetic interference, where it was virtually impossible to use low-frequency instrumentation. Unfortunately, the technique is inapplicable for precision-based physical studies, for example, analysis of the influence of various technological operations on the noise levels in devices. However, the method can be useful for cases where the intensity of the LF noise of the transistor is higher than the noise level of the instrumentation. Such situations are possible, for example, in studies into the influence of long-term external factors, such as temperature, radiation, vibration on the device, as their gradual accumulation leads over time to degradation of the device characteristics. Degradation is always accompanied by the accumulation of various atomic defects; the process leads to an increase rather than a decrease in transistor noise, whose level begins to exceed the noise level of the instrumentation.

REFERENCES

1. **Van der Ziel A.,** Fluctuation phenomena in semi-conductors, Butterworths, London, 1959.
2. **Naryshkin A. K., Vrachev A. S.,** Theory of low-frequency noise, Energiya Publishing, Moscow, 1972 (in Russian).
3. **Buckingham M. J.,** Noise in electronic devices and systems, Halsted Press Publisher, Sydney, 1983.
4. **Dyakonova N. V., Levinshtein M. E., Rumyantsev S. L.,** Priroda ob'emnogo shuma1/*f* v GaAs and Si (obzor) [The nature of volume 1/*f* noise in GaAs and Si (review), Physics & Technology of Semiconductors. 25 (12) (1991) 2065–2104 (in Russian).
5. **Zhao M., Liu X., Wei K., Jin Zh.,** Low frequency noise measurements as a characterization tool for reliability assessment in AlGaN/GaN High-Electron-Mobility Transistors (HEMTs), Proc. 2015 IEEE 11th Int. Conf. Power Electron. Drive Syst., 9–12 June 2015, Sydney, NSW, Australia (2015) 181–183.
6. **Dobrov V. A., Kozlovski V. V., Mescheryakov A. V., et al.,** Effect of electron irradiation with an energy of 0.9 MeV on the *I–V* characteristics and low-frequency noise in 4H–SiC pin diodes, Semiconduct. 53 (4) (2019) 545–551.
7. **Fleetwood D. M.,** Total-ionizing-dose effects, border traps, and 1/*f* noise in emerging MOS technologies, IEEE Trans. Nucl. Sci. 67 (7) (2020) 1216–1240.
8. **Ott H. W.,** Noise reduction techniques in electronic systems, Second edition, Wiley Interscience, New York, 1988.
9. **Levin B. R.,** Teoretycheskiye osnovy statisticheskoy radionekhniki [The theoretical basis of statistical radio engineering], Sovetskoye Radio Publishing, Moscow, 1966 (in Russian).
10. **Kovalev A. N.,** Tranzistory na osnove poluprovodnikovykh geterostruktur [Transistors based on semiconductor heterostructures], MISIS Publishing, Moscow, 2011 (in Russian).
11. **Andrade M. G. C., Bergamim L. F. D. O., Baptista J. B., et al.,** Low-frequency noise investigation of AlGaN/GaN high-electron-mobility transistors, Solid-State Electron. 183 (Sept) (2021) 108050.
12. **Choi Y. J., Lee J. H., An S. J., Im K. S.,** Low-frequency noise behavior of AlGaN/GaN HEMTs with different Al compositions, Crystals. 10 (9) (2020) 830.
13. **Sai P., Jorudas J., Dub M., et al.,** Low frequency noise and trap density in GaN/AlGaN field effect transistors, Appl. Phys. Lett. 115 (18) (2019) 183501.

СПИСОК ЛИТЕРАТУРЫ

1. **Ван дер Зил А.** Флуктуационные явления в полупроводниках. М.: Изд-во иностр. лит., 1961. 232 с.
2. **Нарышкин А. К., Врачев А. С.** Теория низкочастотных шумов. М.: Энергия, 1972. 152 с.
3. **Букингем М.** Шумы в электронных приборах и системах. М.: Мир, 1986. 399 с.
4. **Дьяконова Н. В., Левинштейн М. Е., Румянцев С. Л.** Природа объемного шума 1/*f* в GaAs и Si (обзор) // Физика и техника полупроводников. 1991. Т. 25. № 12. С. 2065–2104.
5. **Zhao M., Liu X., Wei K., Jin Zh.** Low frequency noise measurements as a characterization tool for reliability assessment in AlGaN/GaN High-Electron-Mobility Transistors (HEMTs) // Proceedings of 2015 IEEE 11th International Conference on Power Electronics and Drive Systems. 9–12 June 2015, Sydney, NSW, Australia. 2015. Pp. 181–183.
6. **Добров В. А., Козловский В. В., Мещеряков А. В., Усыченко В. Г., Чернова А. С., Шабунина Е. И., Шмидт Н. М.** Влияние облучения электронами с энергией 0,9 МэВ на вольтамперные характеристики и низкочастотные шумы 4H-SiC pin-диодов // Физика и техника полупроводников. 2019. Т. 53. № 4. С. 555–561.
7. **Fleetwood D. M.** Total-ionizing-dose effects, border traps, and 1/*f* noise in emerging MOS technologies // IEEE Transactions on Nuclear Sciences. 2020. Vol. 67. No. 7. Pp. 1216–1240.
8. **Отт Г.** Методы подавления шумов и помех в электронных системах. М.: Мир, 1979. 318 с.
9. **Левин Б. Р**. Теоретические основы статистической радиотехники. М.: Советское радио, 1966. 728 с.
10. **Ковалев А. Н**. Транзисторы на основе полупроводниковых гетероструктур. М.: Изд-во "МИСИС", 364 .2011 с.

11. **Andrade M. G. C., Bergamim L. F. D. O., Baptista J. B., Nogueira C. R., Silva F. A., Takakura K., Parvais B., Simoen E.** Low-frequency noise investigation of AlGaN/GaN high-electron-mobility transistors // Solid-State Electronics. 2021. Vol. 183. September. P. 108050.

12. **Choi Y. J., Lee J. H., An S. J., Im K. S.** Low-frequency noise behavior of AlGaN/GaN HEMTs with different Al compositions // Crystals. 2020. Vol. 10. No. 9. P. 830.

13. **Sai P., Jorudas J., Dub M., et al.** Low frequency noise and trap density in GaN/AlGaN field effect transistors // Applied Physics Letters. 2019. Vol. 115. No. 18. P. 183501.

THE AUTHORS

USYCHENKO Victor G.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
usychenko@rphf.spbstu.ru
ORCID:009-0002-7673-0918

CHERNOVA Anastasiya S.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
chernova-nastya@mail.ru
ORCID: 0000-0003-1519-0460

СВЕДЕНИЯ ОБ АВТОРАХ

УСЫЧЕНКО Виктор Георгиевич — *доктор физико-математических наук, консультант Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
usychenko@rphf.spbstu.ru
ORCID: 0009-0002-7673-0918

ЧЕРНОВА Анастасия Сергеевна — *ассистентка Высшей школы прикладной физики и космических технологий Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
chernova-nastya@mail.ru
ORCID: 0000-0003-1519-0460

Received 18.01.2024. Approved after reviewing 21.02.2024. Accepted 21.02.2024.
Статья поступила в редакцию 18.01.2024. Одобрена после рецензирования 21.02.2024. Принята 21.02.2024.

PHYSICAL ELECTRONICS

Original article
DOI: https://doi.org/10.18721/JPM.17207

TRANSFORMATION OF THE STRUCTURE OF THIN METAL FILMS UPON ACTIVATION OF THEIR ABILITY TO LOW-VOLTAGE ELECTRON EMISSION

I. S. Bizyaev[1], *P. A. Karaseov*[1], *K. V. Karabeshkin*[2], *P. G. Gabdullin*[1], *A. V. Arkhipov*[1]✉

[1] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia;

[2] Joint-Stock Company "Research and Production Enterprise "ELAR", St. Petersburg, Russia

✉ arkhipov@rphf.spbstu.ru

Abstract. This work continues the studies of low-threshold field electron emission from thin (6–10 nm) films of refractory metals (Mo or Zr) deposited on flat Si substrates. Now, we have investigated the changes in the films' morphology induced by thermo- and electroforming procedures and by the extraction of emission current. In SEM images of the samples taken after emission experiments, we observed the signs of solid-state dewetting (agglomeration) of the films, presumably caused by ion bombardment. This hypothesis was verified by SRIM simulations of the effect of ions on Mo-film/Si-substrate structure, as well as by an experiment at a HVEE-500 ion implanter.

Keywords: thin films, low-field electron emission, film dewetting, ion bombardment

Funding: The reported study was funded by Russian Science Foundation (Grant No. 23-29-10027; https://rscf.ru/project/23-29-10027/), and by St. Petersburg Science Foundation (Grant No. 23-29-10027).

Citation: Bizyaev I. S., Karaseov P. A., Karabeshkin K. V., Gabdullin P. G., Arkhipov A. V., Transformation of the structure of thin metal films upon activation of their ability to low-voltage electron emission, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 80–93. DOI: https://doi.org/10.18721/JPM.17207

Научная статья
УДК 537.533.2, 538.975, 539.1.06
DOI: https://doi.org/10.18721/JPM.17207

ТРАНСФОРМАЦИЯ СТРУКТУРЫ ТОНКИХ МЕТАЛЛИЧЕСКИХ ПЛЕНОК ПРИ АКТИВИРОВАНИИ ИХ СПОСОБНОСТИ К НИЗКОВОЛЬТНОЙ ЭМИССИИ ЭЛЕКТРОНОВ

И. С. Бизяев[1], *П. А. Карасев*[1], *К. В. Карабешкин*[2]
П. Г. Габдуллин[1], *А. В. Архипов*[1]✉

[1] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;

[2] Акционерное общество «Научно–производственное предприятие «ЭЛАР»», Санкт-Петербург, Россия

✉ arkhipov@rphf.spbstu.ru

Аннотация. Настоящая работа является продолжением исследований свойств низкопороговой автоэлектронной эмиссии из тонких (10 − 6 нм) пленок металлов Mo и Zr, сформированных на плоских подложках Si. Теперь изучались изменения морфологии пленок, вызываемые термополевым активированием и отбором эмиссионного тока. Основным экспериментальным методом была электронная микроскопия. Выдвинута гипотеза, что механизм указанных воздействий можно описать как твердотельный деветтинг (агломерация) покрытия, подвергнутого ионной бомбардировке. Для ее проверки средствами пакета SRIM проведено численное моделирование воздействия ионов на структуру Mo-пленка − Si-подложка, а также поставлен эксперимент с использованием ионного имплантера HVEE-500.

Ключевые слова: тонкая металлическая пленка, низковольтная эмиссия электронов, деветтинг пленки, ионная бомбардировка

Финансирование: Исследование поддержано средствами гранта Российского научного фонда № 23-29-10027 (https://rscf.ru/project/23-29-10027/) и гранта Санкт-Петербургского научного фонда № 23-29-10027.

Ссылка для цитирования: Бизяев И. С., Карасев П. А., Карабешкин К. В., Габдуллин П. Г., Архипов А. В. Трансформация структуры тонких металлических пленок при активировании их способности к низковольтной эмиссии электронов // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 2. С. 80–93. DOI: https://doi.org/10.18721/JPM.17207

Introduction

Many types of nanostructured materials and films are capable of emitting electrons at room temperature in moderate electric fields (several V/μm or less) [1–3]. This is unsurprising in carbon nanotubes, nanowalls, fibers and other similar structures capable of providing significant local geometric amplification of the applied field at the tips and edges. However, structures relatively smooth surfaces that do not contain morphological elements with a high geometric aspect ratio were also found to be capable of low-threshold cold emission [4–11]. In particular, we previously found that this property may be inherent in thin metal films, namely, molybdenum and zirconium films with the thickness ranging from 6 to 10 nm deposited on the surface of naturally oxidized silicon. The results of the experiments shedding light on the problem were presented in [12, 13]. In particular, it was reported that the threshold current of 100 nA/s can be extracted in samples of such films at the flat top of a cylindrical anode 6 mm in diameter (located 0.5–0.6 mm away from the sample) for macroscopic electric fields as low as of 1.8–6.4 V/μm. It was also observed

that the structure of the coating had to be partially transformed from initially solid to that containing nanoislands to stabilize the emission current. Thermal field (TF) treatment consisting of electroforming at elevated temperature was performed for this purpose.

This work focuses on the mechanism behind the transformation of the structure of emissive coatings due to TF treatment and/or prolonged extraction of emission current.

There is much interest in this phenomenon because of the potential applications of metal island films, for instance, in plasmonics and in thermoelectric converters.

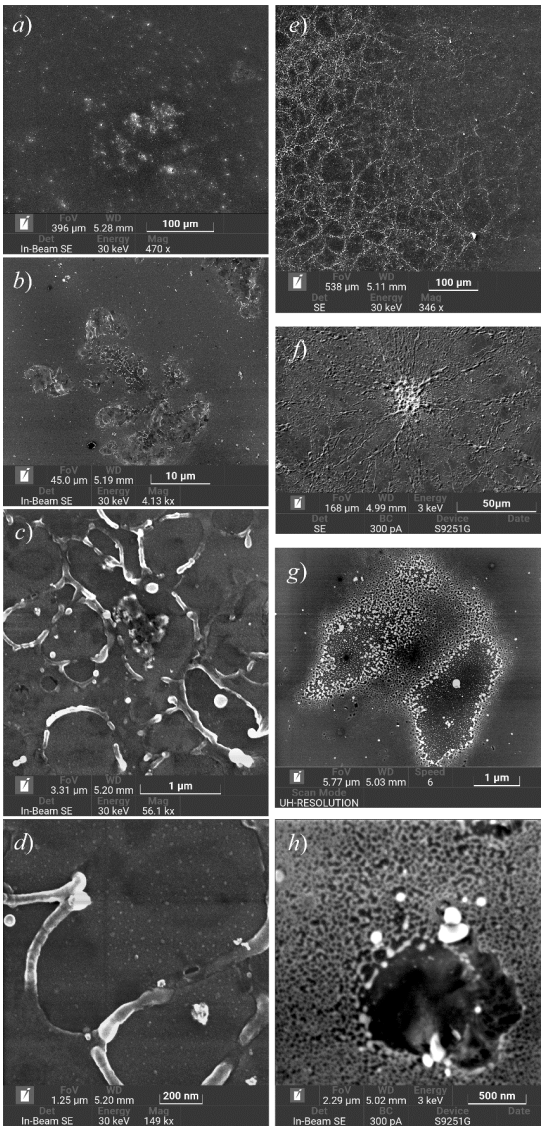

Fig. 1. Surface SEM images for 10 nm thick Mo film samples (*a*–*d*) and 8 nm thick Zr film samples (*e*–*h*) after emission experiments (scale bars correspond to different resolutions)

Experimental procedure

The studies were performed for the same metal film samples as in our earlier works [12, 13]. Magnetron sputtering was used to deposit molybdenum and zirconium films whose thickness ranged from 8 to 10 nm on naturally oxidized plates of KDB–10 grade monocrystalline silicon with *p*-type conductivity. A Mantis HEX system (Mantis Deposition, Thame, UK) was used for deposition; sputtering of the targets was carried out in argon atmosphere at a pressure of $5 \cdot 10^{-3}$ mbar at a rate of about 0.1–1.0 E/s, the temperature of the substrates was 100–150 °C. The thickness of the deposited coatings was determined by a quartz balance.

Before the experiments to determine the capability for emission started, the samples were baked in vacuum for several hours for surface degassing. After that, the samples were subjected to electroforming or TF treatment at temperatures up to 600 °C. The emission characteristics were measured at room temperature and residual pressure of the order of 10^{-9} mbar.

An HVEE 500 kV ion implanter was used in the experiment to simulate the effects of ion irradiation on the coating structure. The sample was irradiated with 10 keV fluorine ions at room temperature and/or at 500 °C.

In all cases, data on the structure of the films were obtained using scanning electron microscopes (SEM) manufactured by Tescan, Czech Republic (Mira, Lyra and Solaris models). The resulting images were processed using the Gwyddion program to eliminate technical defects and increase contrast.

Experimental results and discussion

Microscopic images of the surface. Microscopic images of Mo and Zr films with an average thickness of 8 or 10 nm were obtained using the SEM method. Analysis of the films indicates that the coatings were solid prior to the emission experiments and consisted of grains with characteristic lateral dimensions of the order of tens of nanometers. One of these images (also presented in previous publications [12, 13]) is given below. The surface of the coatings from which the highest emission currents were extracted was considerably transformed after the emission experiments: regions with a modified (damaged) structure appeared, evidently acting as the centers of low-threshold cold electron emission. Different samples exhibited markedly different transformations.

Fig. 1, *a*–*d* shows surface images of the Mo film with the effective thickness of 10 nm yielding the best emission properties: the threshold macroscopic field strength (for a current of 100 nA) was 1.8 V/μm with the extracted stabilized current of the order of tens of μA for tens of hours [13]. Evidently, the main type of defects were circular, crater-like holes ranging in size from fractions to tens of μm, sometimes merging with each other. Comprehensive analysis of high-resolution SEM images (see Fig. 1, *c*,*d*) as well as data from X-ray microanalysis and atomic force microscopy confirmed that the amount of molybdenum inside the craters is reduced tenfold and the metal in these regions is preserved only as separate nanoscale-size islands. The craters are surrounded by elevated rims to which the coating material migrates. Notably, the redistribution of the material did not affect the substrate; its exposed regions remained flat.

The described reconstruction behavior of the metal coating (Fig. 1, *a*–*d*) was achieved for a Mo film sample yielding the best stable emission properties, which is why it was subjected to the longest emission tests with the highest extracted current. Somewhat different types of damage were accumulated in other samples under TF treatment, electroforming and current extraction.

As an example, Fig. 1, *e*–*h* shows SEM images for 8 nm thick Zr film, obtained after emission testing. The overview images (see Fig. 1, *e, f*) show an extensive network of coating regions containing line defects. Their fine structure can be seen in the images with the highest magnification (see Fig. 1, *g*, *h*). The same as in the case above, the process of morphological transformation of the film apparently began with the formation of submicron-sized holes (see Fig. 1,*h*). However, a slightly different picture was subsequently observed for the material from the vacated regions of the surface, composing relatively large (up to 100 nm) circular islands (see Fig. 1, *g*) rather than rims.

Noteworthy structures were also detected on the surfaces of several samples of metal films whose emission properties rapidly degraded during testing. Fig. 2, *a*–*c* shows SEM images of the regions of the 8 nm thick Mo film, Fig. 2, *d*–*f* those of the regions of the Zr film with the same thickness. The structures shown in the overview images *a*, *b*, *d* and *e* consist of several (4–6) relatively wide branches emanating from a common center. As evident from the SEM images, a fraction of the film material moved from the inside of each branch to the surrounding rim. The large-scale images in Fig. 2, *c*, *f* show a network of light lines, probably also corresponding to the line defects on the surface.

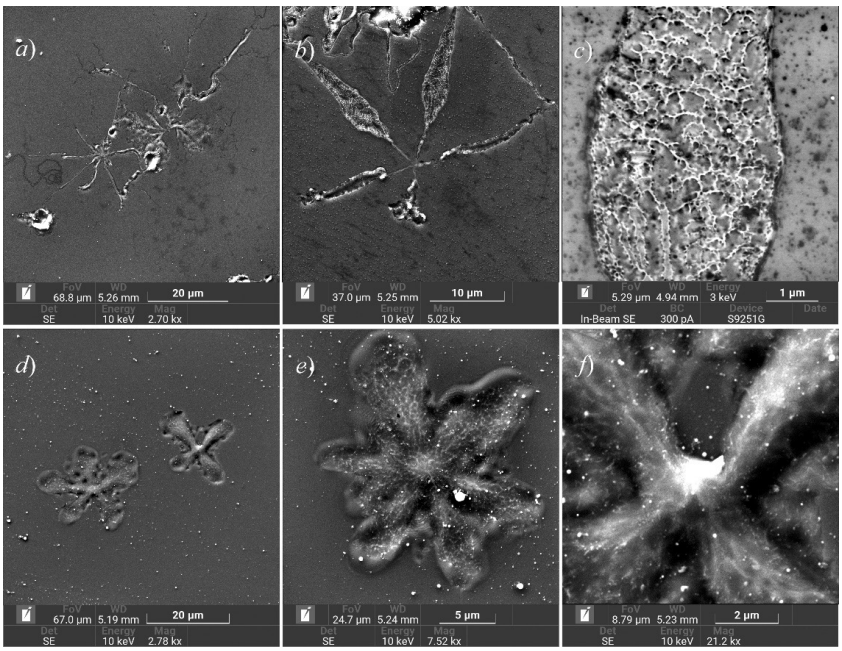

Fig. 2. Branched structures found in SEM images of Mo (*a*–*c*) and Zr (*d*–*e*) coatings (8 nm thick)
Images shown in panels *c* and *f* represent the same regions of the coatings as in *b* and *d*, but at a higher magnification

Thus, we can confirm that regions with modified morphology appeared in the initially solid thin-film samples due to the influence of factors associated with the activation of emission ability. Agglomeration of the films occurred in this case: the amount of coating material in some regions of the surface was significantly reduced due to its lateral transfer to other regions. In some cases, the film structure clearly started to contain islands in some regions (see Fig. 1 *d*, *g*, *h*), which may explain the ability to low-threshold cold electron emission observed in such samples [4, 9, 10].

Mechanism of film agglomeration. The SEM images obtained (see Figs. 1 and 2) combined with the findings from earlier studies using AFM and X-ray microanalysis [12, 13] indicate that the damage in the coatings occurred during emission experiments with metal films largely due to activated material transfer along the surface rather than due to its evaporation or atomization, which could be expected upon rapid local energy release, for example, during electrical breakdown.

The transformation of solid thin metal films into island ones due to lateral transfer of the material is known as solid-state dewetting (the term *agglomeration* is also used in Russian-language literature) [14–16]. When metal is deposited on a dielectric substrate (poorly wetting it), a solid coating can form at initial low temperatures due to low surface mobility of atoms. Such a coating is transformed into an island coating during subsequent thermal annealing, with temperatures significantly below the melting point of the metal generally required for this purpose. The cause of the dewetting phenomenon is the thermodynamic instability of films whose thicknesses range from units to tens of nanometers. This instability is due to the peculiarities of interatomic forces, specifically, the repulsive (short-range) and attractive (long-range) van der Waals forces [17]; or, in another formulation, the forces of diffusion and surface tension.

The general picture (see, for example, [14–18]) is that the dewetting process of a solid coating begins with the appearance of small holes (craters) whose number and size gradually increase. The coating material transferred from the hole forms elevated rims surrounding them. Next, the holes begin to come into contact, and the final coating structure is formed as a result of redistribution of the material in the rims. It can be different and depends on the parameters of the interaction of the atoms in the coating material with each other and with the substrate as well as on the annealing conditions. The formation of both large discrete particles, similar to those shown in Fig. 1,*g*, and branched (also called *web-like*) structures [15, 18]), similar to those shown in Fig. 1, *e*, *f* and Fig. 2, *b*, *d*, was previously described in the literature.

Thus, there is every reason to interpret the appearance of most types of defects in metal films subjected to TF treatment and emission experiments as dewetting occuring under the influence of factors related to extraction of the emission current. In particular, this interpretation allows to explain the correlation of the emission ability observed earlier (in experiments with carbon films [9, 10]) with the presence of a natural oxide layer on silicon substrates: the adhesion of coatings to silicon dioxide is typically lower than their adhesion to pure silicon surface, which can facilitate the transformation of the film structure during activation of their emission ability.

Estimation of dewetting conditions. Thus, an island film is the thermodynamically stable morphological shape for metal coatings up to several tens of nanometers thick deposited on dielectric substrates, since its free energy is lower than the free energy of a solid coating [14–16]. However, solid coatings are transformed into island coatings only at elevated temperatures, since this requires lateral atomic transfer with sufficiently high activation energies. The temperature at which dewetting can be observed depends on the material and thickness of the coating, as well as on the characteristics of the substrate.

Consider gold coatings whose quantitative patterns of dewetting are the best understood. For instance, the activation energy of dewetting was determined in [16] for gold coatings of different thicknesses on silicon nitride substrates. Its values were 1.04 eV for 15 nm thick films and 1.42 eV for 22 nm thick films. The observations were carried out in the temperature range from 300 to 600 °C, which is significantly below the melting point of gold (1064 °C).

The TF treatment aimed at activation of the emission ability of films in our experiments was also carried out at temperatures up to 600 °C. However, we considered metals with much higher melting points (their tabular values are 2623 °C for molybdenum and 1852 °C for zirconium).

Notably, there are scarce data in the literature on dewetting in molybdenum or zirconium films. A significant study [18] observed the process of agglomeration in a 20 nm thick Mo coating deposited on a sapphire substrate. The dewetting of this coating occurred within tens or hundreds of minutes, which approximately corresponds to the duration of the TF treatment in our studies.

However, the samples in [18] were heated to a temperature of 940 °C, significantly higher than the values that we used. This observation is in good agreement with the fact that Mo films exhibited no morphological transformations outside the «projection» of the anode in the region where the only factor influencing the film was the elevated temperature, while the influence of the electric field and the factors associated with extraction of emission current was excluded (this is discussed in our earlier paper [13]). The specific local influence of these factors might have been the reason for the observed local transformation of the film structure and activation of their emission ability.

It was found [19–23] that the dewetting process can be stimulated not only by heating the sample, but also by optical, electronic or ion irradiation (induced by local heating). Estimates of the expected effectiveness of such local contributors are given below for experimental conditions described in [12, 13].

Let us evaluate the temperature of local Joule heat released from the flow of emission current through the emission centers. TF treatment of the samples in [12, 13] was carried out with the extracted emission current $I = 100$ nA $= 10^{-7}$ A. Suppose that as current flows through the emission center, each electron transfers energy of the order $\Delta E = 1$ eV to heat (this assumption is substantiated in [24]). The power of such a heat source is $P = I \cdot \Delta E/e = 10^{-7}$ W $= 100$ NW. An initial estimate of the temperature drop ΔT produced by this source can be obtained by solving the simple problem on the propagation of a stationary heat flow from a local surface source of size d into a thick silicon wafer (with the thermal conductivity $\kappa = 148$ W/(m·K)). The result of the solution is

$$\Delta T \approx \frac{P}{2\pi\kappa d}. \qquad (1)$$

If we select the typical size of the observed craters $d = 1$ μm as the size of the heat release region, we obtain a negligible value for estimating the temperature increment: $\Delta T \approx 10^{-4}$ K. Apparently, the local temperature increment remains insignificant, amounting to $\Delta T \approx 0.01$ K, even for the smallest possible size of the heat release region $d = 10$ nm (the order of size of individual nanostructures detectable in microscopic images in Fig. 1, *d* and *h*).

However, the above evaluation ignores the specifics of nanoscale heat transfer processes, where the contribution from the interface between the media to the total thermal resistance of structures turns out to be predominant [25]. If at least one of the contacting media is not a metal and heat is transferred mainly by phonons, each interface is characterized by the so-called Kapitza resistance R_K [26]. This parameter relates the interfacial temperature drop to the heat flux q: $\Delta T = R_K \cdot q$. In effect, it turns out [25, 27, 28] that the value of the Kapitza resistance for most practically interesting cases lies in a relatively narrow range:

$$R_K = (0.6 - 3.0) \cdot 10^{-8} \text{ m}^2 \cdot \text{K/W}.$$

The maximum estimate of ΔT is obtained taking the value of R_K at the right endpoint of the given range and setting the smallest size of the heat release region: $d = 10$ nm.

Estimating the heat flux across the boundary as $q = P/d^2$, we obtain that

$$\Delta T = R_K \cdot P/d^2 \approx 30 \text{ K}.$$

This value still seems insufficient for explaining the observed transformations in the morphology of the coatings.

The above estimates indicate that local heating of the sample surface induced by the flow of emission current cannot activate the film dewetting process to a sufficient extent.

Let us consider the scenario when the activating factor was the bombardment of the coating regions adjacent to the emission center with ions formed by ionization of residual gas by the emitted electron current. Fig. 3 shows a schematic representation of the experimental device used for TF treatment of samples and measurement of emission current–voltage characteristics.

Let us again set the value of the emission current $I = 100$ nA. The concentration of neutral molecules in the field gap, n, can be estimated based on the residual gas pressure of 10^{-8} Torr (it may be higher than in other parts of the experimental device), which corresponds to $n \approx 3.6 \cdot 10^{14}$ m^{-3}. Generally speaking, the ionization cross section of gas molecules by electron impact depends on

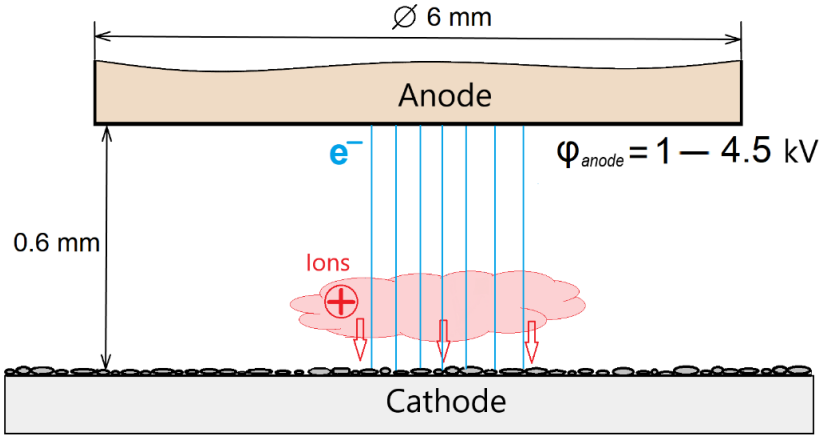

Fig. 3. Schematic of field gap of experimental device
The region colored in pink is where the energy of the emitted electron current (blue lines) reaches the maximum ionization cross section of residual gas and where the largest number of ions (+) bombarding the sample surface are produced (the direction of bombardment is shown by arrows)

the type of gas and on the electron energy of the electrons, varying as they move in the field gap (see Fig. 3). We choose the value of the cross section $\sigma = 10^{-20}$ m^2 as an averaged estimate. The width of the field gap in the experiments was $h = 0.6$ mm. In view of this, the ion current I_i can be written as

$$I_i = I \cdot n\sigma h \approx 2 \cdot 10^{-15} \text{ A}.$$

Next, we estimate the ion beam fluence D for time $t = 1000$ s, corresponding in order of magnitude to the typical duration of the TF treatment procedure. We assume that the ion current is distributed over an area $S = 1$ μm$^2 = 10^{-12}$ m^2 (this is the area of a typical crater in microscopic images, see Fig. 1, *c*, *d* and *h*). We obtain:

$$D = \frac{I_i t}{eS} \approx 10^{18} \text{ m}^{-2} = 10^{14} \text{ cm}^{-2}. \quad (2)$$

This value can be compared with the literature data on ion beam fluences sufficient for initiating dewetting in metal films. For example, a transformation (caused by ion irradiation) of a solid gold film deposited on monocrystalline silicon was observed in [23]; the solid film was transformed into a coating consisting of discrete islands with the lateral size of the order of tens of micrometers. This transformation required irradiation with a fluence in the range $(4.0–9.5) \cdot 10^{13}$ cm^{-2}, which is comparable to estimate (2).

Notably, however, there are several significant differences in the conditions of the compared experiments. The sample considered in [23] was a 1.5 nm thick gold film, while our experiments were carried out with more refractory metals and with coatings of greater thickness. These two differences would be expected to increase the activation energy of atomic migration on the surface. On the other hand, PF_4^+, BF_2^+ and Ta^{4+} ions with sufficiently high energy (0.65 keV/u) were used in [23], typically penetrating deep into the samples. This can weaken the activating effect of radiation on surface migration. The nature of the coating materials seems to be the most significant among these factors; unfortunately, we were unable to uncover any data on the effect of ion irradiation on dewetting of molybdenum and zirconium films in the literature. For this reason, we decided to conduct an additional experiment to directly observe the effect of ion irradiation on one of the film samples whose properties were studied in this paper and earlier in [12, 13].

Effects of ion irradiation: experiment and simulation

Experimental The effect of ion irradiation on the morphology of a Mo film was studied based on the experiment with a sample of 10 nm thick Mo coating on a KDB-10 grade silicon plate.

The peripheral part of the film was used for irradiation and microscopic studies, laying outside the region to which the electric field was applied during the earlier emission testing and therefore not contributing to the emission current.

At the first stage of the experiment, two sections of the sample were exposed to ion irradiation with a fluence (dose) of 10^{12} and 10^{13} cm^{-2} at room temperature. Fluorine ions whose mass is close to atomic masses of atmospheric gases were used. The ion energy was equal to 10 keV, i.e., the minimum value technically achievable for implantation system used. After morphological control of the irradiated and non-irradiated parts of the coating, the sample was reloaded into the implanter and irradiated with the same ions for a second time, but at a dose of 10^{14} cm^{-2} and at a temperature of 500 °C.

Fig. 4 shows images of the sample surface before and after irradiation with doses of 10^{13} and 10^{14} cm^{-2}. The brightest points in the images correspond to defects produced due to prolonged irradiation of the surface by a stopped scanning electron beam.

Even though the maximum fluence was equal to the calculated value (from Eq. (2)), no signs of coating dewetting could be detected in the SEM images. The only transformation observed is in the structure of the grains whose boundaries are practically absent in Fig. 4,*c*, obtained after

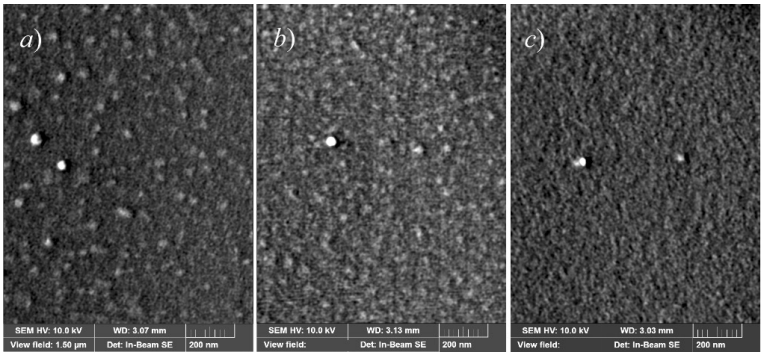

Fig. 4. Surface SEM images for 10 nm thick Mo coating sample obtained at 23 °C before ($D = 0$) (*a*) and after ($D > 0$) single (*b*) and double (*c*) irradiation with 10 keV F^+ ions with different fluences D, cm^{-2}: 10^{13} (*b*); 10^{13} (23 °C) + 10^{14} (500 °C) (*c*)

irradiation with the highest ion fluence, which can be interpreted as a sign of amorphization of the coating.

Coating dewetting at such a fluence value occurred in the experiments with films of more fusible gold [23], even in the case of irradiation at room temperature. On the other hand, the studies in [23] also found that the smoothing of the surface of the gold film at relatively small irradiation doses precedes the period when the film starts to break into islands. This suggests that coating dewetting could also be achieved with a further increase in the dose in our experiment with the implanter.

A significant difference (for the dewetting process) between the ion bombardment accompanying the extraction of the emission current and that in the experiment with the implanter is possibly in the lower average ion energy in the first case, which should lead to a difference in the penetration depth of these ions into the sample.

The maximum ion energy in the emission experiments was determined by the magnitude of the voltage applied to the field gap, which did not exceed 4.5 kV. As for samples with the best emission ability, this voltage did not exceed 1 kV [12, 13]. The average ion energy could be even lower: a wide energy distribution from zero to the value corresponding to the anode potential can be expected for ions produced due to ionization of the residual gas by electron impact (see Fig. 3).

Numerical simulation To confirm the possible role of the voltage factor applied to the field gap, we carried out numerical simulation of the interaction between the ions normally incident on 8 nm thick Mo film on silicon substrate and this structure. The standard SRIM package which is the most widespread for simulations of ion implantation was used. The threshold displacement energies were assumed to be 25 and 15 eV for molybdenum and silicon atoms, respectively. To reduce the statistical error, 15,000 independent cases were simulated. The simulation results are shown in Fig. 5.

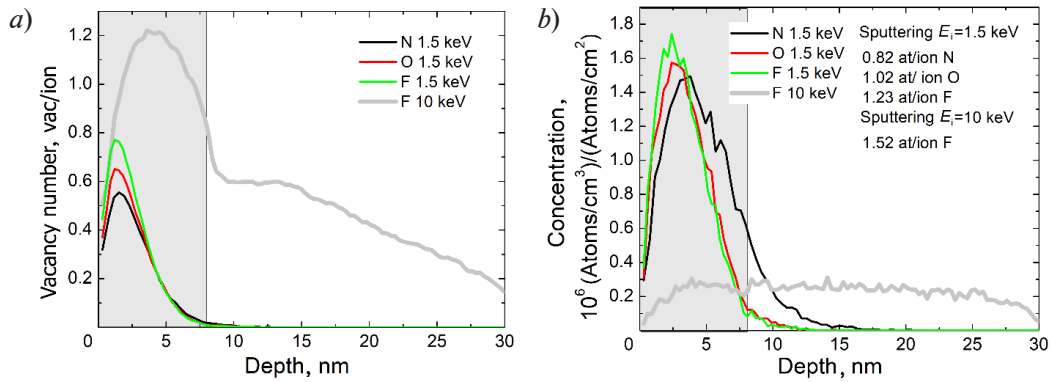

Fig. 5. Calculated distributions of defects produced per ion over target depth (*a*) and concentrations of injected impurity (*b*) in Mo (8 nm)-Si structure during bombardment with different ions (given in the legend)

According to the calculation results, 10 keV ions corresponding to the conditions of the implanter experiment penetrate all the way through the Mo film and a significant number of defects are produced in the surface layer of the silicon substrate. This should adversely affect the coating dewetting process, contributing to mutual diffusion of the film materials, the substrate and the silicon dioxide layer separating them. However, in the case of 1.5 keV ions, it is only the coating that is mostly exposed to them, while the interface with the substrate is virtually unaffected. Therefore, the ion beam bombarding the surface of the samples during TF treatment and emission current extraction in our experiments described in [12, 13] could more effectively stimulate the dewetting process, transferring additional energy to the coating atoms and contributing to an increase in their surface mobility.

The proposed explanation should not be considered as the only possible one. For example, it is known that the action of the electric field itself can increase the surface mobility of atoms and cause agglomeration of films [15]. Studies on electron emission in current flow along thin metal films (see, for example, [29]) interpreted the activating effect of electroforming on coatings precisely as the increase in atomic mobility in a strong electric field. Another potential activating factor is the excitation of electrons in the atomic outer shells by direct action of the electron emission current; this phenomenon of film agglomeration electron irradiation was also described in the literature [19, 20].

Conclusion

The paper analyzes microscopic surface images of thin molybdenum and zirconium film samples, both initially solid and after their thermal field treatment consisting of applying an electric field and extracting an emission current under heating to a predetermined temperature (no higher than 600 °C). The samples subjected to this procedure exhibited (to a greater or lesser extent) cold electron emission in a low-intensity macroscopic electric field (on the order of several V/μm). It was found that the best emission properties of 10 nm thick molybdenum films correlate with a certain type of transformation of their surface during treatment and current extraction, i.e., holes evolving in the film, inside which the coating metal is present as isolated nanoislands. Regions of a different structure appeared during treatment and current extraction in the samples exhibiting the worst emission parameters (in particular, in 8 nm thick molybdenum and zirconium films). The results of image analysis, numerical estimates, simulation and additional experiments indicate that the transformation of the coating structure occurred by the mechanism of solid-state dewetting stimulated by ion bombardment. This opens up the possibilities for developing ion-beam technologies aimed at synthesis of effective emission or thermoelectric structures based on arrays of metal islands.

REFERENCES

1. **Fursey G. N.,** Field emission in vacuum microelectronics, Kluwer Academic–Plenum Publishers, New York, USA, 2005.

2. **Evtukh A., Hartnagel H., Yilmazoglu O., et al.,** Vacuum nanoelectronic devices: Novel electron sources and applications, John Wiley & Sons, Hoboken, USA, 2015.

3. **Egorov N. V., Sheshin E. P.,** Field emission electronics, Springer, Berlin, Germany, 2017.

4. **Eidelman E. D., Arkhipov A. V.,** Field emission from carbon nanostructures: Models and experiment, Phys.–Usp. 63 (7) (2020) 648–667.

5. **Bandurin D. A., Mingels S., Kleshch V. I., et al.,** Field emission spectroscopy evidence for dual-barrier electron tunnelling in nanographite, Appl. Phys. Lett. 106 (23) (2015) 233112.

6. **Fursey G. N., Polyakov M. A., Bagraev N. T., et al.,** Low-threshold field emission from carbon structures, J. Surf. Invest.: X-Ray, Synchrotr. Neutron Tech. 13 (5) (2019) 814–824.

7. **Davidovich M. V., Yafarov R. K.,** Pulsed and static field emission VAC of carbon nanocluster structures: Experiment and its interpretation, Tech. Phys. 64 (8) (2019) 1210–1220.

8. **Streletskiy O. A., Zavidovskiy I. A., Nishchak O. Y., et al.,** Low-threshold field emission cathode based on heat-treated dehydrofluorinated polyvinylidene fluoride, JETP. 135 (6) (2022) 844–852.

9. **Andronov A., Budylina E., Shkitun P., et al.**, Characterization of thin carbon films capable of low-field electron emission, J. Vac. Sci. Technol. B. 36 (2) (2018) 02C108.

10. **Gabdullin P., Zhurkin A., Osipov V., et al.,** Thin carbon films: Correlation between morphology and field-emission capability, Diam. Relat. Mater. 105 (May) (2020) 107805.

11. **Smerdov R., Mustafaev A.,** Novel low-macroscopic-field emission cathodes for electron probe spectroscopy systems, J. Appl. Phys. 134 (11) (2023) 114903.

12. **Bizyaev I. S., Gabdullin P. G., Gnuchev N. M., Arkhipov A. V.,** Low-field electron emission from thin films of metals, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 14 (1) (2021) 105–120.

13. **Bizyaev I., Gabdullin P., Chumak M., et al.,** Low-field electron emission capability of thin films on flat silicon substrates: Experiments with Mo and general model for refractory metals and carbon, Nanomaterials. 11 (12) (2021) 3350.

14. **Thompson C. V.,** Solid-state dewetting of thin films, Annu. Rev. Mater. Res. 42 (1) (2012) 399–434.

15. **Altomare M., Nguyen N. T., Schmuki P.,** Templated dewetting: designing entirely selforganized platforms for photocatalysis, Chem. Sci. 7 (12) (2016) 6865–6886.

16. **Niekiel F., Schweizer P., Kraschewski S. M., et al.,** The process of solid-state dewetting of Au thin films studied by in situ scanning transmission electron microscopy, Acta Mater. 90 (15 May) (2015) 118–132.

17. **Sharma A., Khanna R.,** Pattern formation in unstable thin liquid films, Phys. Rev. Lett. 81 (16) (1998) 3463–3466.

18. **Kovalenko O., Szaby S., Klinger L., Rabkin E.,** Solid state dewetting of polycrystalline Mo film on sapphire, Acta Mater. 139 (15 Oct) (2017) 51–61.

19. **Kojima Y., Kato T.,** Nanoparticle formation in Au thin films by electron-beam-induced dewetting, Nanotechnology. 19 (25) (2008) 255605.

20. **Ruffino F., Grimaldi G.,** Controlled dewetting as fabrication and patterning strategy for metal nanostructures, Phys. Status Solidi A. 212 (8) (2015) 1662–1684.

21. **Lian J. Wang L., Sun X., et al.,** Patterning metallic nanostructures by ion-beam-induced dewetting and Rayleigh instability, Nano Lett. 6 (5) (2006) 1047–1052.

22. **Savio R. L., Repetto L., Guida P., et al.,** Control of the micrometric scale morphology of silicon nanowires through ion irradiation-induced metal dewetting, Solid State Commun. 240 (Aug) (2016) 41–45.

23. **Tuzhilkin M. S., Bespalova P. G., Mishin M. V., et al.,** Formation of Au nanoparticles and features of etching of a Si substrate under irradiation with atomic and molecular ions, Semiconduct. 54 (1) (2020) 137–143.

24. **Arkhipov A. V., Eidelman E. D., Zhurkin A. M., et al.,** Low-field electron emission from carbon cluster films: Combined thermoelectric/hot-electron model of the phenomenon, Fuller. Nanotub. Car. N. 28 (4) (2020) 286–294.

25. **Cahill D. G., Ford W. K., Goodson K. E., et al.,** Nanoscale thermal transport, J. Appl. Phys. 93 (2) (2003) 793–818.

26. **Khalatnikov I. M.,** An introduction to the theory of superfluidity, Perseus Publishing, Cambridge, MA, 2000.

27. **Stoner R. J., Maris H. J.,** Kapitza conductance and heat flow between solids at temperatures from 50 to 300 K, Phys. Rev. B. 48 (22) (1993) 16373–16387.

28. **Meilakhs A. P.,** Nonequilibrium distribution function in the presence of a heat flux at the interface between two crystals, Phys. Solid State. 57 (1) (2015) 148–152.

29. **Fedorovich R. D., Naumovets A. G., Tomchuk P. M.,** Electron and light emission from island metal films and generation of hot electrons in nanoparticles, Phys. Rep. 328 (2–3) (2000) 73–179.

СПИСОК ЛИТЕРАТУРЫ

1. **Fursey G. N.** Field emission in vacuum microelectronics. New York: Kluwer Academic□Plenum Publishers, 2005. 205 p.

2. **Evtukh A., Hartnagel H., Yilmazoglu O., Mimura H., Pavlidis D.** Vacuum nanoelectronic devices: Novel electron sources and applications, Hoboken, USA: John Wiley & Sons, 2015. 453 p.

3. **Шешин Е. П., Егоров Н. В.** Автоэлектронная эмиссия. Принципы и приборы. М.: ИД «Интеллект», 2011. 704 с.

4. **Эйдельман Е. Д., Архипов А. В.** Полевая эмиссия из углеродных наноструктур: модели и эксперимент // Успехи физических наук. 2020. Т. 190. № 7. С. 693–714.

5. **Bandurin D. A., Mingels S., Kleshch V. I., Luetzenkirchen-Hecht D., Mueller G., Obraztsov A. N.** Field emission spectroscopy evidence for dual-barrier electron tunnelling in nanographite // Applied Physics Letters. 2015. Vol. 106. No. 23. P. 233112.

6. **Фурсей Г. Н., Поляков М. А., Баграев Н. Т., Закиров И. И., Нащекин А. В., Бочаров В. Н.** Низкопороговая полевая эмиссия из углеродных структур // Поверхность. Рентгеновские, синхротронные и нейтронные исследования. 2019. № 9. С. 28–39.

7. **Давидович М. В., Яфаров Р. К.** Импульсные и статические автоэмиссионные ВАХ углеродных нанокластерных структур: эксперимент и его интерпретация // Журнал технической физики. 2019. Т. 8 № .89. С. 1293–1282.

8. **Стрелецкий О. А., Завидовский И. А., Нищак О. Ю., Хайдаров А. А., Савченко Н. Ф., Павликов А. В.** Низкопороговый автоэмиссионный катод на основе термически обработанного дегидрофторированного поливинилиденфторида // Журнал экспериментальной и теоретической физики. 2022. Т. 162. № 6. С. 881–891.

9. **Andronov A., Budylina E., Shkitun P., Gabdullin P., Gnuchev N., Kvashenkina O., Arkhipov A.** Characterization of thin carbon films capable of low-field electron emission // Journal of Vacuum Science & Technology B. 2018. Vol. 36. No. 2. P. 02C108.

10. **Gabdullin P., Zhurkin A., Osipov V., Besedina N., Kvashenkina O., Arkhipov A.** Thin carbon films: Correlation between morphology and field-emission capability // Diamond & Related Materials. 2020. Vol. 105. May. P. 107805.

11**. Smerdov R., Mustafaev A.** Novel low-macroscopic-field emission cathodes for electron probe spectroscopy systems // Journal of Applied Physics. 2023. Vol. 134. No. 11. P. 114903.

12. **Бизяев И. С., Габдуллин П. Г., Гнучев Н. М., Архипов А. В.** Низкопороговая полевая эмиссия электронов тонкими металлическими пленками: связь эмиссионной способности с составом, условиями нанесения и морфологией // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2021. Т. 14. № 1. С. 111–127.

13. **Bizyaev I., Gabdullin P., Chumak M., Babyuk V., Davydov S., Osipov V., Kuznetsov A., Kvashenkina O., Arkhipov A.** Low-field electron emission capability of thin films on flat silicon substrates: Experiments with Mo and general model for refractory metals and carbon // Nanomaterials. 2021. Vol. 11. No. 12. P. 3350.

14. **Thompson C. V.** Solid-state dewetting of thin films // Annual Review of Materials Research. 2012. Vol. 42. No. 1. Pp. 399–434.

15. **Altomare M., Nguyen N. T., Schmuki P.** Templated dewetting: designing entirely selforganized platforms for photocatalysis // Chemical Science. 2016. Vol. 7. No. 12. Pp. 6865–6886.

16. **Niekiel F., Schweizer P., Kraschewski S. M., Butz B., Spiecker E.** The process of solid-state dewetting of Au thin films studied by in situ scanning transmission electron microscopy // Acta Materialia. 2015. Vol. 90. 15 May. Pp. 118–132.

17. **Sharma A., Khanna R.** Pattern formation in unstable thin liquid films // Physical Review Letters. 1998. Vol. 81. No. 16. Pp. 3463–3466.

18. **Kovalenko O., Szaby S., Klinger L., Rabkin E.** Solid state dewetting of polycrystalline Mo film on sapphire // Acta Materialia. 2017. Vol. 139. 15 Oct. Pp. 51–61.

19. **Kojima Y., Kato T.** Nanoparticle formation in Au thin films by electron-beam-induced dewetting // Nanotechnology. 2008. Vol. 19. No. 25. P. 255605.

20. **Ruffino F., Grimaldi G.** Controlled dewetting as fabrication and patterning strategy for metal nanostructures // Physica Status Solidi A. 2015. Vol. 212. No. 8. Pp.1662–1684.

21. **Lian J. Wang L., Sun X., Yu Q., Ewing R. C**. Patterning metallic nanostructures by ion-beam-induced dewetting and Rayleigh instability // Nano Letters. 2006. Vol. 6. No. 5. Pp. 1047–1052.

22. **Savio R. L., Repetto L., Guida P., Angeli E., Firpo G., Volpe A., Ierardi V., Valbusa U.** Control of the micrometric scale morphology of silicon nanowires through ion irradiation-induced metal dewetting // Solid State Communications. 2016. Vol. 240. August. Pp. 41–45.

23. **Тужилкин М. С., Беспалова П. Г., Мишин М. В., Колесников И. Е., Карабешкин К. В., Карасев П. А., Титов А. И.** Формирование наночастиц Au и особенности травления подложки Si после облучения атомарными и молекулярными ионами // Физика и техника полупроводников. 2020. Т. 54. № 1. С. 90–96.

24. **Arkhipov A. V., Eidelman E. D., Zhurkin A. M., Osipov V. S., Gabdullin P. G.** Low-field electron emission from carbon cluster films: Combined thermoelectric/hot-electron model of the phenomenon // Fullerenes, Nanotubes and Carbon Nanostructures. 2020. Vol. 28. No. 4. Pp. 286–294.

25. **Cahill D. G., Ford W. K., Goodson K. E., Mahan G. D., Majumdar A., Maris H. J., Merlin R., Phillpot S. R.** Nanoscale thermal transport // Journal of Applied Physics. 2003. Vol. 93. No. 2. Pp. 793–818.

26. **Халатников И. М.** Введение в теорию сверхтекучести. М.: Наука, 1965. 158 с.

27. **Stoner R. J., Maris H. J.** Kapitza conductance and heat flow between solids at temperatures from 50 to 300 K // Physical Review B. 1993. Vol. 48. No. 22. Pp. 16373–16387.

28. **Мейлахс А.П.** Неравновесная функция распределения при тепловом потоке вблизи границы двух кристаллов // Физика твердого тела. 2015. Т. 57. № 1. С. 140–144.

29. **Fedorovich R. D., Naumovets A. G., Tomchuk P. M.** Electron and light emission from island metal films and generation of hot electrons in nanoparticles // Physics Reports. 2000. Vol. 328. No. 2–3. Pp. 73–179.

THE AUTHORS

BIZYAEV Ivan S.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
ivanbiziaev@yandex.com
ORCID: 0009-0005-7265-1173

KARASEOV Platon A.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
platon.karaseov@spbstu.ru
ORCID: 0000-0003-2511-0188

KARABESHKIN Konstantin V.
Joint-Stock Company "Research and Production Enterprise «ELAR»"
68 Torez Ave., St. Petersburg, 194223, Russia
yanikolaus@yandex.ru
ORCID: 0000-0003-1770-1877

GABDULLIN Pavel G.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
gabdullin_pg@spbstu.ru
ORCID: 0000-0002-2519-2577

ARKHIPOV Alexander V.
Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
arkhipov@rphf.spbstu.ru
ORCID: 0000-0002-3321-7797

СВЕДЕНИЯ ОБ АВТОРАХ

БИЗЯЕВ Иван Сергеевич – *аспирант Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого, научный сотрудник лаборатории «Самоорганизующиеся высокотемпературные наноструктуры» Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
ivanbiziaev@yandex.com
ORCID: 0009-0005-7265-1173

КАРАСЕВ Платон Александрович – *доктор физико-математических наук, профессор Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
platon.karaseov@spbstu.ru
ORCID: 0000-0003-2511-0188

КАРАБЕШКИН Константин Валерьевич – *кандидат физико-математических наук, инженер Акционерного общества "Научно–производственное предприятие «ЭЛАР»».*
194223, Россия. г. Санкт-Петербург, пр. Тореза, 68, лит. А
yanikolaus@yandex.ru
ORCID: 0000-0003-1770-1877

ГАБДУЛЛИН Павел Гарифович – *кандидат физико-математических наук, доцент Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
gabdullin_pg@spbstu.ru
ORCID: 0000-0002-2519-2577

АРХИПОВ Александр Викторович – *доктор физико-математических наук, профессор Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
arkhipov@rphf.spbstu.ru
ORCID: 0000-0002-3321-7797

Received 21.01.2024. Approved after reviewing 11.02.2024. Accepted 11.02.2024.
Статья поступила в редакцию 21.01.2024. Одобрена после рецензирования 11.02.2024. Принята 11.02.2024.

Original article
DOI: https://doi.org/10.18721/JPM.17208

ANALYTICAL ELECTRIC POTENTIALS FOR SIMULATION OF THE MULTIPOLE RADIOFREQUENCY ION FUNNELS WITH NONLINEAR PROFILES

A. S. Berdnikov[1]✉, *S.V. Masyukevich*[1], *T. V. Pomozov*[1], *Yu. I. Hasin*[1], *K. V. Solovyev*[2,1]

[1] Institute for Analytical Instrumentation, RAS, St. Petersburg, Russia;

[2] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ asberd@yandex.ru

Abstract. The article considers analytical expressions for electric field potentials that correspond to radio-frequency funnels with quadratic, cubic and biquadratic profiles. The funnel electrodes are a sequence of diaphragms with a circular or multipole structure and apertures of variable size. The latter varies along the transport channel according to the desired law. The resulting expressions can be useful for fast, high-quality simulation of radiofrequency devices designed to convey ion beams with simultaneous focusing.

Keywords: Laplace equation, periodic electrodes, analytical electric field, ion guide, ion funnel

Funding: The reported study was carried out at Institute for Analytical Instrumentation of Russian Academy of Sciences (St. Petersburg, Russia) within the framework of the theme FFZM-2022-0009 (State registration No. 122040600002-3) of the State Assignment of the Ministry of Science and Higher Education of the Russian Federation, No. 075-01157-23-00 dated 29.12.2022.

Citation: Berdnikov A. S., Masyukevich S. V., Pomozov T. V., Hasin Yu. I., Solovyev K. V., Analytical electric potentials for simulation of the multipole radiofrequency ion funnels with nonlinear profiles, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 94–119. DOI: https://doi.org/10.18721/JPM.17208

Научная статья
УДК 537.21, 517.958
DOI: https://doi.org/10.18721/JPM.17208

АНАЛИТИЧЕСКИЕ ПОТЕНЦИАЛЫ ЭЛЕКТРИЧЕСКИХ ПОЛЕЙ ДЛЯ МОДЕЛИРОВАНИЯ МУЛЬТИПОЛЬНЫХ РАДИОЧАСТОТНЫХ ИОННЫХ ВОРОНОК С НЕЛИНЕЙНЫМ ПРОФИЛЕМ

А. С. Бердников[1]✉, *С. В. Масюкевич*[1], *Т. В. Помозов*[1], *Ю. И. Хасин*[1], *К. В. Соловьев*[2,1]

[1] Институт аналитического приборостроения РАН, Санкт-Петербург, Россия;

[2] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉asberd@yandex.ru

Аннотация. В статье рассматриваются аналитические выражения для потенциалов электрических полей, которые соответствуют радиочастотным воронкам с квадратичным, кубическим и биквадратным профилями. Электроды воронок представляют собой последовательность диафрагм с круговой или мультипольной структурой и апертурами переменного размера, который изменяется вдоль канала транспортировки по заданному закону. Полученные выражения могут быть полезны для быстрого качественного моделирования радиочастотных устройств, предназначенных для транспортировки и фокусировки ионов.

Ключевые слова: уравнение Лапласа, периодические электроды, аналитическое электрическое поле, ионное транспортирующее устройство

Финансирование: Работа выполнена в Институте аналитического приборостроения Российской академии наук (Санкт-Петербург) в рамках темы FFZM-2022-0009 (номер гос. регистрации 122040600002-3) государственного задания Министерства науки и высшего образования Российской Федерации № 075-01157-23-00 от 29.12.2022.

Ссылка для цитирования: Бердников А. С., Масюкевич С. В., Помозов Т. В., Хасин Ю. И., Соловьев К. В. Аналитические потенциалы электрических полей для моделирования мультипольных радиочастотных ионных воронок с нелинейным профилем // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 2. С. 94–119. DOI: https://doi.org/10.18721/JPM.

Introduction

The present study is a direct continuation of the analysis in [1], considering analytical expressions for the potentials of electric fields in cylindrical transporting channels [2, 3] (in fact, in SRIG-type cylindrical radio frequency traps, first proposed in [4, 5] and studied in detail in [2, 6–9]), and conical ion funnels [10–16]. However, when designing ion-optical devices, it is convenient to use RF focusing funnels whose profile is not conical (this is shown, for example, in [13]), as well as segmented quadrupole or multipole electrodes (see, for example, [33]).

The analytical electrical potentials considered in this paper are intended for modeling ion funnels with quadratic, cubic and biquadratic profiles, and, more generally, for modeling SRIG–type transport systems with circular, non-circular and multipole apertures, as well as with profiles described by polynomials of the appropriate degree.

These analytical solutions of the three-dimensional Laplace equation are of separate interest, since they can be used to solve the corresponding problems of mathematical physics.

RF traps with a cylindrical transport channel

Traps with circular electrodes. In the simplest case, RF traps and SRIG-type (Stacked Ring Ion Guides) transport channels are a sequence of circular apertures with RF voltages applied to them (in particular, sequences of pulses) with a phase shift π between adjacent apertures [2–10]. Assuming that the condition of quasi-static electric field is satisfied, the high-frequency potential of this field can be expressed as a function of time setting the law for voltage variation at the electrodes, which is multiplied by the potential of the electrostatic field corresponding to DC voltages at the electrodes.

The assumption about the quasi-static nature of a high-frequency electric field is valid when the time of the characteristic change in electrical voltages at the electrodes significantly exceeds the time of propagation of electromagnetic perturbation within the device. Typical sizes of electrode configurations used in the designs of ion-optical systems are no more than a few tens of centimeters (as a rule, much smaller). Taking into account the equality of the speed of light and the speed of propagation of an electromagnetic perturbation, this assumption is obviously satisfied for the frequencies of electrical voltages commonly used in ion-optical devices (from hundreds of kilohertz to several megahertz, rarely several gigahertz).

Under such conditions, the electromagnetic perturbation propagates almost instantly, and the high-frequency electric potential, which varies in time and space, can indeed be represented as a product of a function of time and electrostatic potential depending on spatial coordinates. A time-dependent multiplier describes a temporal change in electrical voltages, and a coordinate-dependent potential corresponds to constant voltages at the electrodes, which change synchronously and proportionally over time.

Thus, an electric field is created within the framework of the model under consideration, which changes synchronously and proportionally at each point in space. Although such a step is, in fact, a disregard for electrodynamic effects, i.e., the accompanying electromagnetic wave, it is quite acceptable if the frequency of voltages applied to the electrodes is not too high (in the sense indicated above).

In a typical case, the signs and amplitudes of the voltages applied to individual apertures of a cylindrical trap or a cylindrical transport channel alternate according to the rule

$$+U_R, -U_R, +U_R, -U_R, +U_R, -U_R, \ \ldots \tag{1}$$

Due to the symmetry of the geometric shape of the electrodes and the antisymmetry of the voltages applied to the electrodes, the plane along which the electric potential is identically zero is located in the middle between adjacent electrodes. This makes it possible to add fictitious zero-voltage apertures without distorting the electric field and thus consider periodic sequences of voltages that alternate according to another rule:

$$+U_R, 0, -U_R, 0, +U_R, 0, -U_R, 0, +U_R, 0, \ -U_R, \ldots \tag{2}$$

If the electrode configuration and the analytical expression for the electric potential with voltages (2) applied to the apertures are shifted by one step along the z axis, a solution is obtained for the voltages at the apertures organized by the rule

$$0, +U_R, 0, -U_R, 0, +U_R, 0, -U_R, 0, +U_R, 0, \ldots \tag{3}$$

Fig. 1 shows a SRIG-type cylindrical RF trap with circular apertures.

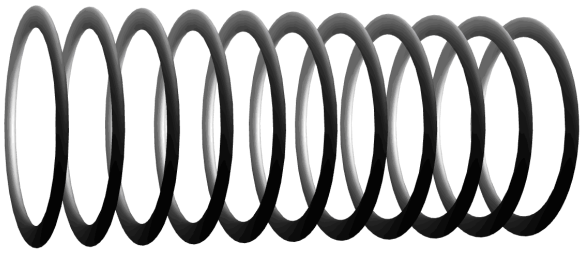

Fig. 1. Structure of periodic ring electrodes of a cylindrical RF trap or a cylindrical SRIG-type transport channel

It is known [2] that the electrostatic potential of a system of electrodes with the periodic voltages applied to them corresponding to sequence (1) is described with good accuracy by the following expression away from the edges of the electrodes, in the vicinity of the axis of the transport channel:

$$U(z,r)=\frac{U_R}{I_0(\lambda R)}\cos(\lambda z+\varphi)I_0(\lambda r), \tag{4}$$

where z, r are the axial and radial coordinates ($r=\sqrt{x^2+y^2}$); U_R are the static potentials applied to the apertures; I_0 is the modified zero-order Bessel function [17–19]; λ is a parameter of the geometric scale, $\lambda=\pi/L$ (L is the distance between adjacent apertures); R is the radius of circular apertures; φ is the parameter determining the shift of the aperture sequence relative to the origin.

Equipotential lines and a three-dimensional graph of the electric potential given by Eq. (4) in normalized coordinates (z, r) are shown in Fig. 2.

Potential (4) includes the sum of two linearly independent potentials with constant coefficients:

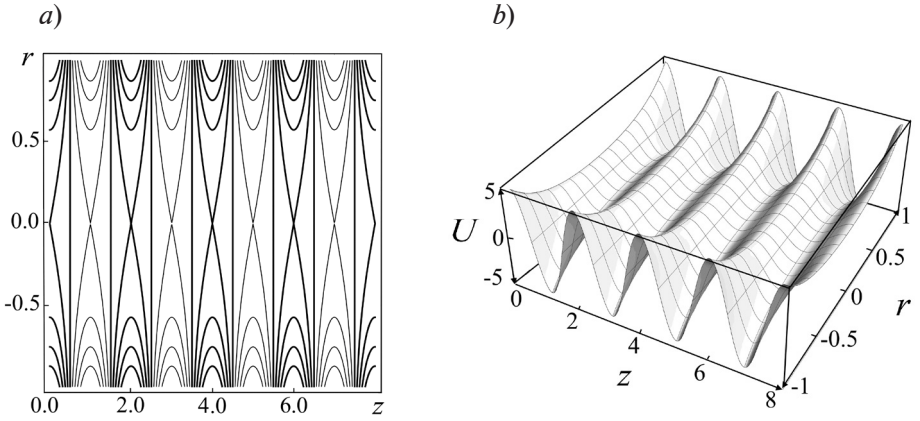

Fig. 2. Field lines for static electric potential (5) in (r,z) plane in normalized coordinates for SRIG-type RF trap:
a corresponds to equipotential lines on the plane, *b* to the three-dimensional graph

$$U_C(z,r)=\frac{U_R}{I_0(\lambda R)}\cos(\lambda z)I_0(\lambda r), \tag{5}$$

$$U_S(z,r)=\frac{U_R}{I_0(\lambda R)}\sin(\lambda z)I_0(\lambda r), \tag{6}$$

Potential (5) behaves as $U_0\cos(\lambda z)$ on the axis of symmetry $r=0$, and potential (6) behaves as $U_0\sin(\lambda z)$, where U_0 is the amplitude of spatial oscillations of the corresponding electrostatic potential on the axis of the system, $U_0=U_R/I_0(\lambda R)$. Electrostatic potentials (5), (6) prove to be useful auxiliary tools for constructing analytical models of the electric field for more complex electrode configurations.

Functions (5) can be used to describe the electrostatic field of a periodic sequence of circular apertures with the voltages corresponding to periodic sequence (2), and functions (6) can be used to describe apertures with the voltages corresponding to periodic sequence (3). To achieve this, however, it is necessary to use a geometric scale $\lambda=\pi/(2L)$.

Indeed, the voltage chains (2) and (3) are obtained from voltage chain (1) by inserting additional zero-voltage apertures between the initial apertures (as noted earlier, this operation does not distort the electric field for a sequence of geometrically identical electrodes with antisymmetric periodic voltages of form (1)). The distance between adjacent apertures becomes equal to $L/2$ on the initial scale of distances, so the geometric scale λ should be recalculated.

A practical example of using potentials (5), (6) in combination with circuits for supplying electrical voltages (2) and (3) are analytical electric fields used for modeling ion-optical devices [8, 9, 21–32].

Traps with quadrupole segmentation of electrodes. Segmented quadrupole apertures can be used instead of circular apertures (Fig. 3). Namely, circular apertures are divided into four segments, and then electric voltages are applied to the resulting electrodes in such a way that opposite-polarity electric voltages appear at adjacent electrodes. Such a device is the first stage of a two-stage transport channel, considered, in particular, in [33], using numerical simulation rather than analytical models of the electric potential.

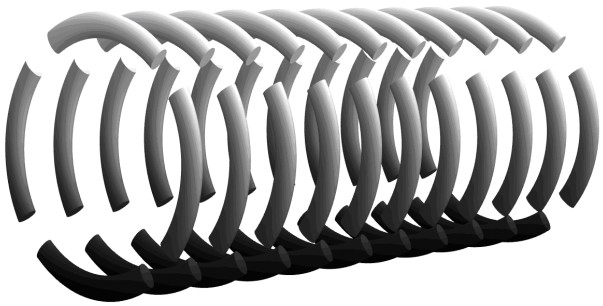

Fig. 3. Structure of periodic segmented quadrupole electrodes of cylindrical RF trap

A good approximation for the electrode configuration shown in Fig. 3, in the axial region sufficiently away from the edges of the circular electrodes, is expressed as electrostatic potentials taking the form

$$U_C^{(q)}(x,y,z) = \frac{U_R^{(q)}}{I_2(\lambda R)} \frac{x^2 - y^2}{r^2} \cos(\lambda z)\, I_2(\lambda r), \tag{7}$$

$$U_S^{(q)}(x,y,z) = \frac{U_R^{(q)}}{I_2(\lambda R)} \frac{x^2 - y^2}{r^2} \sin(\lambda z)\, I_2(\lambda r), \tag{8}$$

where $r = \sqrt{x^2 + y^2}$ (as before); $\pm U_R^{(q)}$ are the electrical voltages with alternating signs, applied to individual multipole segments of a thin quadrupole-segmented aperture; I_2 is the modified second-order Bessel function [17–19].

The estimate $I_2(r) \sim r^2/8$ is valid for the function I_2 with $r \approx 0$. Because the limit of the expression $I_2(\lambda r)/r^2$ is finite for $r \to 0$, potential (7) behaves as follows near the symmetry axis $r = 0$:

$$U_C^{(q)} \sim U_0^{(q)} (x^2 - y^2) \cos(\lambda z),$$

and potential (8) behaves as

$$U_S^{(q)} \sim U_0^{(q)} (x^2 - y^2) \sin(\lambda z),$$

where $U_0^{(q)}$ is the spatial oscillation amplitude of the corresponding quadrupole field component near the system axis,

$$U_0^{(q)} = \lambda^2 U_R^{(q)} / 8 I_2(\lambda R).$$

Good approximations of the electric field in the axial region for the electrode configuration, rotated relative to the axis of symmetry by 45, are analytical functions of the form

$$U_C^{(r)}(x,y,z)=\frac{U_R^{(r)}}{I_2(\lambda R)}\frac{xy}{r^2}\cos(\lambda z)\,I_2(\lambda r), \tag{9}$$

$$U_S^{(r)}(x,y,z)=\frac{U_R^{(r)}}{I_2(\lambda R)}\frac{xy}{r^2}\sin(\lambda z)\,I_2(\lambda r). \tag{10}$$

Potential (9) behaves as $\sim U_0^{(r)} xy\cos(\lambda z)$ near the symmetry axis $r = 0$, and potential (10) behaves as $\sim U_0^{(r)} xy\sin(\lambda z)$.

Analytical solutions (7)–(10) can be combined to obtain an analytical solution for a segmented quadrupole trap with alternating orientations of adjacent electrodes (Fig. 4).

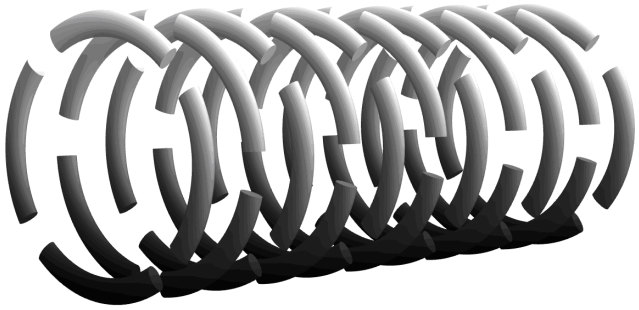

Fig. 4. Structure of periodic segmented quadrupole electrodes of cylindrical RF trap with quasi-octupole placement symmetry of even and odd electrodes

We use voltage sequence (2) for circular apertures with traditional quadrupole segmentation (see Fig. 3). The electrostatic potential of such a system is described with good accuracy by analytical expression (7), using the geometric scale $\lambda = \pi/(2L)$. The electrostatic potential for a sequence of voltages (3) and circular apertures with rotated quadrupole segmentation is described with good accuracy by analytical expression (8) with the geometric scale $\lambda = \pi/(2L)$. Now let us combine these two solutions to obtain an analytical expression of the electric potential for the configuration of the circular apertures in Fig. 4.

The shape of circular electrodes in the corresponding section is of no particular consequence for zero potentials in a first approximation. The circular segmented quadrupole electrodes for even voltage positions (2) generating electrostatic potential (7) within the volume of the trap can be replaced with rotated segmented quadrupole electrodes. Similarly, the rotated circular segmented quadrupole electrodes for odd voltage positions (3) generating electrostatic potential (7) within the volume of the trap can be replaced with conventional segmented quadrupole electrodes. In accordance with the principle of superposition of electric fields, summation of voltages for a fixed geometric structure of electrodes involves summation of the corresponding electric fields. This yields additional analytical solutions describing the electrostatic field for a trap with rotated segmented quadrupole electrodes (see Fig. 4):

$$U^{(a)}(x,y,z)=\pm U_C^{(q)}(x,y,z)\pm U_S^{(r)}(x,y,z), \tag{11}$$

$$U^{(b)}(x,y,z)=\pm U_C^{(r)}(x,y,z)\pm U_S^{(q)}(x,y,z), \tag{12}$$

where the functions $U_C^{(q)}$, $U_S^{(q)}$, $U_C^{(r)}$, $U_S^{(r)}$ are defined by expressions (7)–(10), and the geometric scale parameter is selected in accordance with equality $\lambda = \pi/(2L)$.

Note. The given scheme for combining two analytical solutions can be used with different electrical voltages, different radii of circular apertures, different electrode segmentation schemes (see the next section) and even with different shapes of non-circular apertures for even and odd positions. However, such exotic solutions are of theoretical rather than practical interest.

Traps with multipole segmentation of electrodes. In general, a multipole of arbitrary order n can be used instead of quadrupole segmentation of circular electrodes. For this purpose, circular apertures are divided into $2n$ identical segments, and then electric voltages are applied to the resulting electrodes in such a way that electric voltages of opposite polarity appear at adjacent electrodes. A good approximation in the axial region sufficiently away from the edges of the circular electrodes (depending on the rotation of segmented multipole electrodes relative to the symmetry axis) is expressed as electrostatic potentials taking the form

$$U_C^{(q,n)}(x,y,z) = \frac{U_R}{I_n(\lambda R)} r^n \cos\left(n \operatorname{Arg}(x,y)\right)\cos(\lambda z) I_n(\lambda r)/r^n, \tag{13}$$

$$U_S^{(q,n)}(x,y,z) = \frac{U_R}{I_n(\lambda R)} r^n \cos\left(n \operatorname{Arg}(x,y)\right)\sin(\lambda z) I_n(\lambda r)/r^n, \tag{14}$$

$$U_C^{(r,n)}(x,y,z) = \frac{U_R}{I_n(\lambda R)} r^n \sin\left(n \operatorname{Arg}(x,y)\right)\cos(\lambda z) I_n(\lambda r)/r^n, \tag{15}$$

$$U_S^{(r,n)}(x,y,z) = \frac{U_R}{I_n(\lambda R)} r^n \sin\left(n \operatorname{Arg}(x,y)\right)\sin(\lambda z) I_n(\lambda r)/r^n, \tag{16}$$

where $\operatorname{Arg}(x,y)$ is the argument of the complex number $x + iy$; $\lambda = \pi/L$ is the parameter of the geometric scale; I_n is the modified Bessel function of nth order [17–19], n is the order of multipole segmentation of circular apertures of a cylindrical trap.

If $r \approx 0$, the function $I_n(r)$ behaves as

$$I_n(r) \sim 2^{-n} r^n / n!,$$

and if $r \to \infty$, it behaves as

$$I_n(r) \sim \exp(r)/\sqrt{2\pi r},$$

In view of this, potentials (13)–(16) behave as follows at $r \approx 0$ (i.e., near the axis of symmetry):

$$U_C^{(q,n)} \sim U_0 r^n \cos\left(n \operatorname{Arg}(x,y)\right)\cos(\lambda z), \tag{17}$$

$$U_S^{(q,n)} \sim U_0 r^n \cos\left(n \operatorname{Arg}(x,y)\right)\sin(\lambda z), \tag{18}$$

$$U_C^{(r,n)} \sim U_0 r^n \sin\left(n \operatorname{Arg}(x,y)\right)\cos(\lambda z), \tag{19}$$

$$U_S^{(r,n)} \sim U_0 r^n \sin\left(n \operatorname{Arg}(x,y)\right)\sin(\lambda z), \tag{20}$$

where U_0 is the spatial oscillation amplitude of the corresponding multipole component of the electric potential near the axis of the system,

$$U_0 = U_R\left(2^{-n}\lambda^n\right)/\left(n! I_n(\lambda R)\right).$$

It is easy to verify that functions (13)–(16) do indeed satisfy the Laplace equation (as well as the asymptotic behavior of functions at $r \to 0$ and $r \to \infty$) by direct substitution using the program from [38]. Eqs. (5), (6) and (7)–(10) are special cases of analytical expressions (13)–(16). Linear combinations with constant coefficients composed of axisymmetric potentials (5), (6) and multipole potentials (7)–(10) and (13)–(16) with appropriate weights can be used to analytically describe the electric fields of RF traps and SRIG-type transport devices with non-circular apertures.

As in the case of circular segmented quadrupoleapertures discussed in the previous section, segmented multipole solutions can be combined and analytical solutions can be obtained for segmented multipole apertures with alternating rotation angles:

$$U^{(a,n)}(x,y,z) = \pm U_C^{(q,n)}(x,y,z) \pm U_S^{(r,n)}(x,y,z), \tag{21}$$

$$U^{(b,n)}(x,y,z) = \pm U_C^{(r,n)}(x,y,z) \pm U_S^{(q,n)}(x,y,z). \tag{22}$$

The functions $U_C^{(q,n)}$, $U_S^{(q,n)}$, $U_C^{(r,n)}$, $U_S^{(r,n)}$ in these expressions are defined by Eqs. (13)–(16), while $\lambda = \pi/(2L)$.

Two-dimensional multipole multipliers. The expressions in Eqs. (13)–(20)

$$Q_n(x,y) = r^n \cos\left(n \operatorname{Arg}(x,y)\right), \tag{23}$$

$$R_n(x,y) = r^n \sin\left(n \operatorname{Arg}(x,y)\right), \tag{24}$$

are homogeneous harmonic polynomials satisfying the two-dimensional Laplace equation. The Euler homogeneity property of degree n for the functions $Q_n(x,y)$ and $R_n(x,y)$follows directly from Eqs. (23), (24):

$$\text{for } \forall\lambda > 0, \forall x,\ y: Q_n(\lambda x, \lambda y) = \lambda^n Q_n(x,y),\ R_n(\lambda x, \lambda y) = \lambda^n R_n(x,y),$$

and harmonicity is verified by direct substitution of expressions (23), (24) into the Laplace equation.

The sequence of two-dimensional homogeneous harmonic polynomials $Q_n(x,y)$ и $R_n(x,y)$, calculated by Eqs. (23), (24) and ordered by increasing homogeneity degree (multipole order), has the form:

$$1, x, y, x^2 - y^2, 2xy, x^3 - 3xy^2, 3x^2y - y^3, \ldots .$$

The multipliers $Q_n(x,y)$ and $R_n(x,y)$ provide multipole symmetry with respect to the symmetry axis OZ for the potentials under consideration. Expressions (23), (24) obey the following recurrence relations:

$$Q_0(x,y) = 1,\ R_0(x,y) = 0, \tag{25}$$

$$Q_{n+1}(x,y) = xQ_n(x,y) - yR_n(x,y), \tag{26}$$

$$R_{n+1}(x,y) = yQ_n(x,y) + xR_n(x,y). \tag{27}$$

General expressions for functions defined using recurrence relations (25)–(27) are Eqs. (23), (24). This either follows from the uniqueness of the functions that must be calculated in accordance with recurrence relations (25)–(27), or is easily proved by induction, or becomes obvious after writing recurrence relations (25)–(27) in complex form:

$$Q_0(x,y) + iR_0(x,y) = 1,$$

$$Q_n(x,y) + iR_n(x,y) = (x+iy)\left(Q_{n-1}(x,y) + iR_{n-1}(x,y)\right),$$

$$Q_n(x,y) + iR_n(x,y) = (x+iy)^n .$$

The last equality means that

$$Q_n = \operatorname{Re}(x+iy)^n,\ R_n = \operatorname{Im}(x+iy)^n,$$

then, the trigonometric notation of complex numbers can be used to obtain formulas (23), (24).

The polynomial nature of functions (23), (24) as well as the homogeneity of these polynomials (each of them is a sum of homogeneous monomials of degree n, which have the form $c_k x^k y^{n-k}$, where $k = 0, 1, ..., n$) automatically follow from recurrence relations (25)–(27). The fact that these polynomials satisfy the two-dimensional Laplace equation can be obtained without using the general formulas (23), (24), as the Cauchy–Riemann relations are satisfied for the functions $Q_n(x,y)$ and $R_n(x, y)$, obeying recurrence relations (25)–(27):

$$\frac{\partial Q_n(x,y)}{\partial x} = \frac{\partial R_n(x,y)}{\partial y}, \quad \frac{\partial Q_n(x,y)}{\partial y} = -\frac{\partial R_n(x,y)}{\partial x}. \tag{28}$$

Property (28) is proved by induction using auxiliary equalities that follow from recurrence relations (25)–(27):

$$\frac{\partial Q_{n+1}}{\partial x} - \frac{\partial R_{n+1}}{\partial y} = Q_n + x\frac{\partial Q_n}{\partial x} - y\frac{\partial R_n}{\partial x} - Q_n - y\frac{\partial Q_n}{\partial y} - x\frac{\partial R_n}{\partial y} = x\left(\frac{\partial Q_n}{\partial x} - \frac{\partial R_n}{\partial y}\right) - y\left(\frac{\partial Q_n}{\partial y} + \frac{\partial R_n}{\partial x}\right),$$

$$\frac{\partial Q_{n+1}}{\partial y} + \frac{\partial R_{n+1}}{\partial x} = x\frac{\partial Q_n}{\partial y} - R_n - y\frac{\partial R_n}{\partial y} + y\frac{\partial Q_n}{\partial x} + R_n + x\frac{\partial R_n}{\partial x} = x\left(\frac{\partial Q_n}{\partial y} + \frac{\partial R_n}{\partial x}\right) + y\left(\frac{\partial Q_n}{\partial x} - \frac{\partial R_n}{\partial y}\right).$$

RF funnels with conical transport channel

Funnels with circular electrodes. The classical version of an RF ion funnel is a sequence of circular apertures with decreasing radii, varying by the law for the radii of the circular cone cut by planes located in the plane of the aperture and perpendicular to the axis of the cone (Fig. 5).

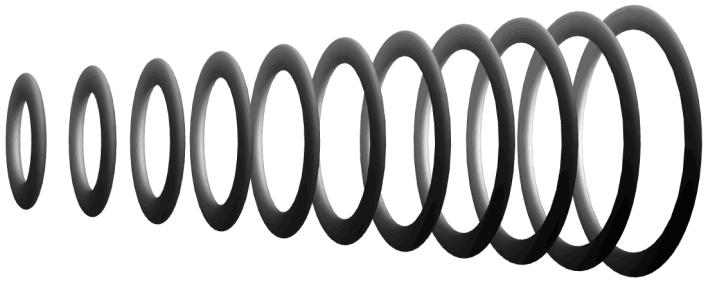

Fig. 5. Structure of periodic ring electrodes of conical RF funnel

The behavior of the axisymmetric electric potential $V(z,r)$ is to be expected for such a geometric configuration of the electrodes on the axis of the system as the sum of potentials

$$V_C \sim U_0 z\cos(\lambda z) \text{ and } V_S \sim U_0 z\sin(\lambda z),$$

where U_0 is the scaling factor for the linearly increasing spatial oscillation amplitude of the corresponding electric potential on the axis of the system.

The following analytical formulas were proposed in [1] for conical RF funnels with circular apertures:

$$V_C(z,r) = U_0\left[z\cos(\lambda z)I_0(\lambda r) + r\sin(\lambda z)I_1(\lambda r)\right], \tag{29}$$

$$V_S(z,r) = U_0\left[z\sin(\lambda z)I_0(\lambda r) - r\cos(\lambda z)I_1(\lambda r)\right], \tag{30}$$

where $\lambda = \pi/L$ is the geometric scale, I_1 is a modified first-order Bessel function [17–19].

If $r \approx 0$, the estimate $I_1(r) \approx r/2$ is valid for the function I_1. These analytical solutions behave as follows on the z axis of the system:

$$V_C \sim U_0 z\cos(\lambda z),\ V_S \sim U_0 z\sin(\lambda z).$$

Equipotential lines and a three-dimensional graph of the electric potential given by Eq. (29) in normalized coordinates (z, r) are shown in Fig. 6. Function (29) turns into function (30) when substituting $z \to z + \pi/2\lambda$, so their graphs merge into each other with a shift along the OZ axis.

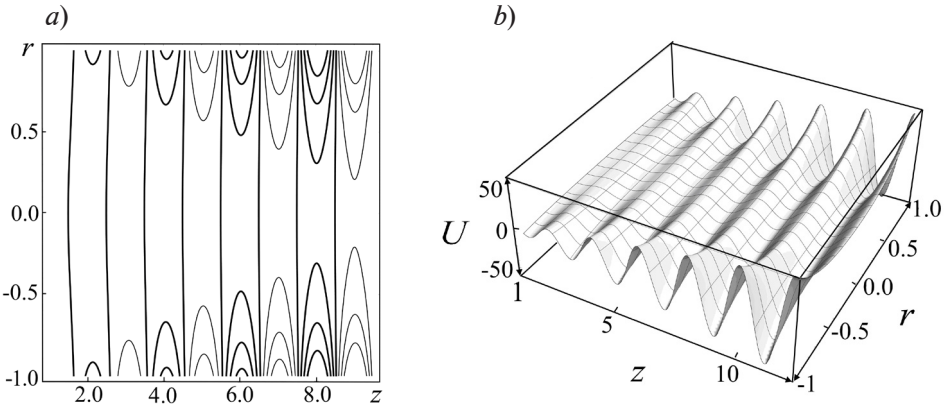

Fig. 6. Field lines for static electric potential (29) in (z,r) plane in normalized coordinates for SRIG-type RF trap:
a corresponds to equipotential lines on the plane, *b* to the three-dimensional graph

Solutions (29), (30) are obtained by differentiation of solutions (5), (6) with respect to the parameter λ after the substitution

$$U_0 = U_R / I_0(\lambda R),$$

where U_0 is the amplitude of spatial oscillations of the electrostatic potential on the axis of the SRIG-type RF trap.

Funnels with segmented quadrupole electrodes. Similarly, differentiating with respect to the parameter λ yields analytical solutions for conical segmented quadrupole funnels from Eqs. (7)–(10) (by analogy with Fig. 3):

$$V_C^{(q)}(x,y,z) = U_0^{(q)}(x^2 - y^2)\left[z\cos(\lambda z)\frac{8I_2(\lambda r)}{\lambda^2 r^2} + \right.$$
$$\left. +4r^2 \sin(\lambda z)\frac{\lambda r\left(I_1(\lambda r) + I_3(\lambda r)\right) - 4I_2(\lambda r)}{\lambda^3 r^2}\right], \quad (31)$$

$$V_S^{(q)}(x,y,z) = U_0^{(q)}(x^2 - y^2)\left[z\sin(\lambda z)\frac{8I_2(\lambda r)}{\lambda^2 r^2} - \right.$$
$$\left. -4r^2 \cos(\lambda z)\frac{\lambda r\left(I_1(\lambda r) + I_3(\lambda r)\right) - 4I_2(\lambda r)}{\lambda^3 r^2}\right]. \quad (32)$$

where I_3 is a modified third-order Bessel function [17–19].

The estimate $I_3(r) \sim r^3/48$ is valid for the function I_3 with $r \approx 0$.

A device such as the second stage of a two-stage transport channel was considered, in particular, in [33] but using numerical simulation, rather than analytical models of the electric potential.

If $r = 0$, i.e., near the symmetry axis, potential (31) behaves as follows:

$$V_C^{(q)} \sim U_0^{(q)}(x^2 - y^2) z\cos(\lambda z),$$

and potential (32) behaves as

$$V_S^{(q)} \sim U_0^{(q)}(x^2 - y^2) z\sin(\lambda z).$$

Similarly, analytical solutions for the corresponding static electric potentials for the configuration of segmented quadrupole circular electrodes rotated 45° relative to the symmetry axis are functions of the form

$$V_C^{(r)}(x,y,z) = U_0^{(r)} xy\left[z\cos(\lambda z)\frac{8I_2(\lambda r)}{\lambda^2 r^2} + 4r\sin(\lambda z)\frac{\lambda r\left(I_1(\lambda r)+I_3(\lambda r)\right)-4I_2(\lambda r)}{\lambda^3 r^3}\right], \quad (33)$$

$$V_S^{(r)}(x,y,z) = U_0^{(r)} xy\left[z\sin(\lambda z)\frac{8I_2(\lambda r)}{\lambda^2 r^2} - 4r\cos(\lambda z)\frac{\lambda r\left(I_1(\lambda r)+I_3(\lambda r)\right)-4I_2(\lambda r)}{\lambda^3 r^3}\right], \quad (34)$$

If $r = 0$, i.e., near the symmetry axis, potential (14) behaves as:

$$V_C^{(r)} \sim U_0^{(r)} xyz\cos(\lambda z),$$

and potential (15) behaves as

$$V_S^{(r)} \sim U_0^{(r)} xyz\sin(\lambda z).$$

As in the case of SRIG-type segmented quadrupole RF traps, by analogy with Fig. 4, combining quadrupole potentials with different rotations of the electrodes relative to the symmetry axis yields analytical models of electric potentials for conical funnels with different rotations of composite electrodes relative to the funnel axis for even and odd apertures:

$$V^{(a)}(x,y,z) = \pm V_C^{(q)}(x,y,z) \pm V_S^{(r)}(x,y,z), \quad (35)$$

$$V^{(b)}(x,y,z) = \pm V_C^{(r)}(x,y,z) \pm V_S^{(q)}(x,y,z). \quad (36)$$

Funnels with multipole-segmented electrodes. In general, analytical solutions for a conical RF trap with segmented multipole electrodes are obtained by differentiating Eqs. (13)–(16) with respect to the parameter λ after substitution

$$U_0 = U_R / I_0(\lambda R),$$

where U_0 is the spatial oscillation amplitude of the corresponding multipole component of the electrostatic potential on the axis of the RF trap:

$$V_C^{(q,n)}(x,y,z) = U_0^{(q)} r^n \cos\left(n\,\mathrm{Arg}(x,y)\right)\left[z\cos(\lambda z) f_0(\lambda r) + r\sin(\lambda z) f_1(\lambda r)\right], \quad (37)$$

$$V_S^{(q,n)}(x,y,z) = U_0^{(q)} r^n \cos\left(n\,\mathrm{Arg}(x,y)\right)\left[z\sin(\lambda z) f_0(\lambda r) - r\cos(\lambda z) f_1(\lambda r)\right], \quad (38)$$

$$V_C^{(r,n)}(x,y,z) = U_0^{(r)} r^n \sin\left(n\,\mathrm{Arg}(x,y)\right)\left[z\cos(\lambda z) f_0(\lambda r) + r\sin(\lambda z) f_1(\lambda r)\right], \quad (39)$$

$$V_S^{(r,n)}(x,y,z) = U_0^{(r)} r^n \sin\left(n\,\mathrm{Arg}(x,y)\right)\left[z\sin(\lambda z) f_0(\lambda r) - r\cos(\lambda z) f_1(\lambda r)\right], \quad (40)$$

where

$$f_0(\rho) = \frac{2^n n! I_n(\rho)}{\rho^n}, \quad (41)$$

$$f_1(\rho) = \frac{2^{n-1} n!}{\rho^{n+1}}\left[\rho\left(I_{n-1}(\rho)+I_{n+1}(\rho)\right) - 2nI_n(\rho)\right]. \quad (42)$$

Expressions (41), (42) have no singularities at zero, and the following equalities are fulfilled for $\rho \approx 0$:

$$f_0(\rho) \approx 1 + \frac{1}{4(n+1)}\rho^2,$$

$$f_1(\rho) \approx \frac{1}{2(n+1)}\rho + \frac{1}{8(n+1)(n+2)}\rho^3.$$

By analogy with SRIG-type segmented multipole RF traps with rotated electrodes for even and odd apertures (see Fig. 4), it is possible to combine multipole potentials with different rotations of multipole configurations, allowing to obtain analytical models of electric potentials for the corresponding conical funnels:

$$V^{(a,n)}(x,y,z) = \pm V_C^{(q,n)}(x,y,z) \pm V_S^{(r,n)}(x,y,z), \tag{43}$$

$$V^{(b,n)}(x,y,z) = \pm V_C^{(r,n)}(x,y,z) \pm V_S^{(q,n)}(x,y,z). \tag{44}$$

Radio frequency funnels with polynomial channel profile

Circular electrodes. It was found in [13] that RF funnels with circular apertures and quadratic profile of the transport channel (Fig. 7) can provide additional advantages compared to conventional conical funnels. Analytical expressions for axisymmetric electric potentials that follow the dependences

$$W_C \sim U_0 z^2 \cos(\lambda z) \text{ and } W_S \sim U_0 z^2 \sin(\lambda z),$$

on the z axis can be obtained by differentiating expressions (29), (30) with respect to the parameter λ.

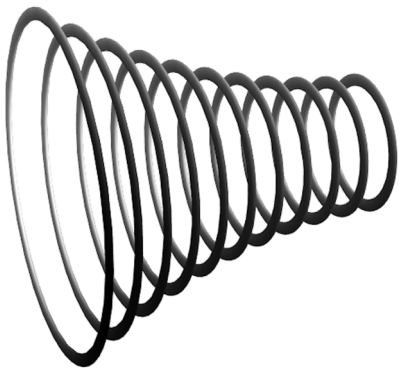

Fig. 7. Structure of periodic circular electrodes of RF funnel with quadratic profile

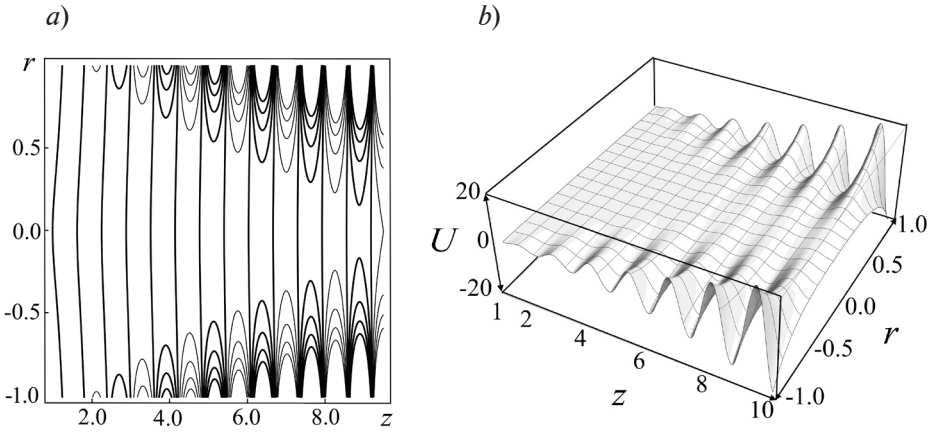

Fig. 8. Field lines for static electric potential (45) in (z,r) plane in normalized coordinates for SRIG-type radio frequency trap:
a corresponds to equipotential lines on the plane, *b* to the three-dimensional graph

This gives the following formulas:

$$W_C(z,r) = U_0\left[z^2\cos(\lambda z)I_0(\lambda r) + 2rz\sin(\lambda z)I_1(\lambda r) - \frac{r^2}{2}\cos(\lambda z)\left(I_0(\lambda r) + I_2(\lambda r)\right)\right], \quad (45)$$

$$W_S(z,r) = U_0\left[z^2\sin(\lambda z)I_0(\lambda r) - 2rz\cos(\lambda z)I_1(\lambda r) - \frac{r^2}{2}\sin(\lambda z)\left(I_0(\lambda r) + I_2(\lambda r)\right)\right]. \quad (46)$$

Equipotential lines and a three-dimensional graph of the electric potential given by Eq. (45) in normalized coordinates (z, r) are shown in Fig. 8. As in the case of funnels with a conical profile, function (45) turns into function (46) upon substituting $z \to z + \pi/2\lambda$, so their graphs merge into each other with a shift along the OZ axis.

Similar analytical expressions are obtained for axisymmetric electric potentials obeying the following dependences on the z axis:

$$G_C \sim U_0 z^3\cos(\lambda z),\ G_S \sim U_0 z^3\sin(\lambda z),\ H_C \sim U_0 z^4\cos(\lambda z),\ H_S \sim U_0 z^4\sin(\lambda z),$$

namely,

$$G_C(z,r) = U_0\left[z^3\cos(\lambda z)I_0(\lambda r) + 3rz^2\sin(\lambda z)I_1(\lambda r) - \right.$$
$$\left. - \frac{3r^2 z}{2}\cos(\lambda z)\left(I_0(\lambda r) + I_2(\lambda r)\right) - \frac{r^3}{4}\sin(\lambda z)\left(3I_1(\lambda r) + I_3(\lambda r)\right)\right], \quad (47)$$

$$G_S(z,r) = U_0\left[z^3\sin(\lambda z)I_0(\lambda r) - 3rz^2\cos(\lambda z)I_1(\lambda r) - \right.$$
$$\left. - \frac{3r^2 z}{2}\sin(\lambda z)\left(I_0(\lambda r) + I_2(\lambda r)\right) + \frac{r^3}{4}\sin(\lambda z)\left(3I_1(\lambda r) + I_3(\lambda r)\right)\right], \quad (48)$$

$$H_C(z,r) = U_0\left[z^4\cos(\lambda z)I_0(\lambda r) + 4rz^2\sin(\lambda z)I_1(\lambda r) - \right.$$
$$-3r^2z^2\cos(\lambda z)\left(I_0(\lambda r) + I_2(\lambda r)\right) - r^3 z\sin(\lambda z)\left(3I_1(\lambda r) + I_3(\lambda r)\right) +$$
$$\left. + \frac{r^4}{8}\cos(\lambda z)\left(3I_0(\lambda r) + 4I_2(\lambda r) + I_4(\lambda r)\right)\right], \quad (49)$$

$$H_S(z,r) = U_0\left[z^4\sin(\lambda z)I_0(\lambda r) - 4rz^2\cos(\lambda z)I_1(\lambda r) - \right.$$
$$-3r^2z^2\sin(\lambda z)\left(I_0(\lambda r) + I_2(\lambda r)\right) + r^3 z\cos(\lambda z)\left(3I_1(\lambda r) + I_3(\lambda r)\right) +$$
$$\left. + \frac{r^4}{8}\sin(\lambda z)\left(3I_0(\lambda r) + 4I_2(\lambda r) + I_4(\lambda r)\right)\right]. \quad (50)$$

Segmented quadrupole and multipole electrodes. Differentiation of expressions (37)–(40) with respect to the parameter λ yields analytical expressions for the electric potentials of RF funnels with multipole segmentation (or with quadrupole segmentation at $n = 2$) of circular apertures, a nonlinear (quadratic, cubic and biquadratic) profile of the transport channel and the corresponding rotation of multipole segments relative to the symmetry axis:

$$W_C^{(q,n)}(z,r) = U_0 r^n\cos\left(n\,\mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^2\cos(\lambda z)f_0(\lambda r) + 2rz\sin(\lambda z)f_1(\lambda r) - r^2\cos(\lambda z)f_2(\lambda r)\right], \quad (51)$$

$$W_S^{(q,n)}(z,r) = U_0 r^n\cos\left(n\,\mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^2\sin(\lambda z)f_0(\lambda r) - 2rz\cos(\lambda z)f_1(\lambda r) - r^2\sin(\lambda z)f_2(\lambda r)\right], \quad (52)$$

$$W_C^{(r,n)}(z,r)=U_0 r^n \sin\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^2\cos(\lambda z)f_0(\lambda r)+2rz\sin(\lambda z)f_1(\lambda r)-r^2\cos(\lambda z)f_2(\lambda r)\right], \tag{53}$$

$$W_S^{(r,n)}(z,r)=U_0 r^n \sin\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^2\sin(\lambda z)f_0(\lambda r)-2rz\cos(\lambda z)f_1(\lambda r)-r^2\sin(\lambda z)f_2(\lambda r)\right], \tag{54}$$

$$G_C^{(q,n)}(z,r)=U_0 r^n \cos\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^3\cos(\lambda z)f_0(\lambda r)+3z^2 r\sin(\lambda z)f_1(\lambda r)-\right.$$
$$\left.-3zr^2\cos(\lambda z)f_2(\lambda r)-r^3 f_3(\lambda r)\sin(\lambda z)\right], \tag{55}$$

$$G_S^{(q,n)}(z,r)=U_0 r^n \cos\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^3\sin(\lambda z)f_0(\lambda r)-3z^2 r\cos(\lambda z)f_1(\lambda r)-\right.$$
$$\left.-3zr^2\sin(\lambda z)f_2(\lambda r)+r^3 f_3(\lambda r)\cos(\lambda z)\right], \tag{56}$$

$$G_C^{(r,n)}(z,r)=U_0 r^n \sin\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^3\cos(\lambda z)f_0(\lambda r)+3z^2 r\sin(\lambda z)f_1(\lambda r)-\right.$$
$$\left.-3zr^2\cos(\lambda z)f_2(\lambda r)-r^3 f_3(\lambda r)\sin(\lambda z)\right], \tag{57}$$

$$G_S^{(r,n)}(z,r)=U_0 r^n \sin\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^3\sin(\lambda z)f_0(\lambda r)-3z^2 r\cos(\lambda z)f_1(\lambda r)-\right.$$
$$\left.-3zr^2\sin(\lambda z)f_2(\lambda r)+r^3 f_3(\lambda r)\cos(\lambda z)\right], \tag{58}$$

$$H_C^{(q,n)}(z,r)=U_0 r^n \cos\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^4\cos(\lambda z)f_0(\lambda r)+4z^3 r\sin(\lambda z)f_1(\lambda r)-\right.$$
$$-6z^2r^2\cos(\lambda z)f_2(\lambda r)-4zr^3 f_3(\lambda r)\sin(\lambda z)+$$
$$\left.+r^4 f_4(\lambda r)\cos(\lambda z)\right], \tag{59}$$

$$H_S^{(q,n)}(z,r)=U_0 r^n \cos\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^4\sin(\lambda z)f_0(\lambda r)-4z^3 r\cos(\lambda z)f_1(\lambda r)-\right.$$
$$-6z^2r^2\sin(\lambda z)f_2(\lambda r)+4zr^3 f_3(\lambda r)\cos(\lambda z)+$$
$$\left.+r^4 f_4(\lambda r)\sin(\lambda z)\right], \tag{60}$$

$$H_C^{(r,n)}(z,r)=U_0 r^n \sin\left(n\ \mathrm{Arg}(x,y)\right)\times$$
$$\times\left[z^4\cos(\lambda z)f_0(\lambda r)+4z^3 r\sin(\lambda z)f_1(\lambda r)-\right.$$
$$-6z^2r^2\cos(\lambda z)f_2(\lambda r)-4zr^3 f_3(\lambda r)\sin(\lambda z)+$$
$$\left.+r^4 f_4(\lambda r)\cos(\lambda z)\right], \tag{61}$$

$$
\begin{aligned}
H_S^{(r,n)}(z,r) &= U_0 r^n \sin\left(n\ \mathrm{Arg}(x,y)\right)\times \\
&\times\Big[z^4\sin(\lambda z) f_0(\lambda r) - 4z^3 r\cos(\lambda z) f_1(\lambda r) - \\
&-6z^2 r^2 \sin(\lambda z) f_2(\lambda r) + 4z r^3 f_3(\lambda r)\cos(\lambda z) + \\
&+r^4 f_4(\lambda r)\sin(\lambda z)\Big],
\end{aligned}
\tag{62}
$$

where

$$f_0(\rho) = \frac{2^n n! I_n(\rho)}{\rho^n},$$

$$f_1(\rho) = \frac{2^{n-1} n!}{\rho^{n+1}}\Big[\rho\left(I_{n-1}(\rho) + I_{n+1}(\rho)\right) - 2nI_n(\rho)\Big],$$

$$
\begin{aligned}
f_2(\rho) = \frac{2^{n-2} n!}{\rho^{n+2}}\Big[&\rho^2\left(I_{n-2}(\rho) + 2I_n(\rho) + I_{n+2}(\rho)\right) - \\
&-\rho\left(I_{n-1}(\rho) + I_{n+1}(\rho)\right) + \\
&+4n(n+1) I_n(\rho)\Big],
\end{aligned}
$$

$$
\begin{aligned}
f_3(\rho) = \frac{2^{n-3} n!}{\rho^{n+3}}\Big[&\rho^3\left(I_{n-3}(\rho) + 3I_{n-1}(\rho) + 3I_{n+1}(\rho) + I_{n+3}(\rho)\right) - \\
&-6n\rho^2\left(I_{n-2}(\rho) + 2I_n(\rho) + I_{n+2}(\rho)\right) + \\
&+12n(n+1)\rho\left(I_{n-1}(\rho) + I_{n+1}(\rho)\right) - \\
&-8n(n+1)(n+2) I_n(\rho)\Big],
\end{aligned}
$$

$$
\begin{aligned}
f_4(\rho) = \frac{2^{n-4} n!}{\rho^{n+4}}\Big[&\rho^4\left(I_{n-4}(\rho) + 4I_{n-2}(\rho) + 6I_n(\rho) + 4I_{n+2}(\rho) + I_{n+4}(\rho)\right) - \\
&-8n\rho^3\left(I_{n-3}(\rho) + 3I_{n-1}(\rho) + 3I_{n+1}(\rho) + I_{n+3}(\rho)\right) + \\
&+24n(n+1)\rho^2\left(I_{n-2}(\rho) + 2I_n(\rho) + I_{n+2}(\rho)\right) - \\
&-32n(n+1)(n+2)\rho\left(I_{n-1}(\rho) + I_{n+1}(\rho)\right) + \\
&+16n(n+1)(n+2)(n+3) I_n(\rho)\Big].
\end{aligned}
$$

The functions $f_0(\rho)$, $f_1(\rho)$, $f_2(\rho)$, $f_3(\rho)$, $f_4(\rho)$ have no singularities at zero; the following approximate equalities are satisfied for $\rho \approx 0$:

$$f_0(\rho) \approx 1 + \frac{1}{4(n+1)}\rho^2,$$

$$f_1(\rho) \approx \frac{1}{2(n+1)}\rho + \frac{1}{8(n+1)(n+2)}\rho^3,$$

$$f_2(\rho) \approx \frac{1}{2(n+1)} + \frac{1}{8(n+1)(n+2)}\rho^2,$$

$$f_3(\rho) \approx \frac{3}{4(n+1)(n+2)}\rho + \frac{5}{16(n+1)(n+2)(n+3)}\rho^3,$$

$$f_4(\rho) \approx \frac{3}{4(n+1)(n+2)} + \frac{15}{16(n+1)(n+2)(n+3)}\rho^2.$$

It is also possible to combine multipole potentials with different rotations of multipole configurations to obtain analytical models of electric potentials for the corresponding funnels with nonlinear profiles:

$$W^{(a,n)}(x,y,z) = \pm W_C^{(q,n)}(x,y,z) \pm W_S^{(r,n)}(x,y,z),$$

$$W^{(b,n)}(x,y,z) = \pm W_C^{(r,n)}(x,y,z) \pm W_S^{(q,n)}(x,y,z),$$

$$G^{(a,n)}(x,y,z) = \pm G_C^{(q,n)}(x,y,z) \pm G_S^{(r,n)}(x,y,z),$$

$$G^{(b,n)}(x,y,z) = \pm G_C^{(r,n)}(x,y,z) \pm G_S^{(q,n)}(x,y,z),$$

$$H^{(a,n)}(x,y,z) = \pm H_C^{(q,n)}(x,y,z) \pm H_S^{(r,n)}(x,y,z),$$

$$H^{(b,n)}(x,y,z) = \pm H_C^{(r,n)}(x,y,z) \pm H_S^{(q,n)}(x,y,z).$$

Estimation of approximation accuracy

As we already discussed, the above analytical expressions do not accurately describe the required electric fields near the edges of the electrodes. It is required to obtain an estimate the distance from the edges of the electrodes sufficient for guaranteeing the obtained analytical expressions.

Consider, for example, the axisymmetric electric potential given by Eq. (6). In the case when the distance between adjacent apertures is equal to L, the radius of the aperture is equal to R, and the voltages at the apertures are set according to Eq. (1), leading to the analytical expression

$$U_S(z,r) = \frac{U_R}{I_0(\pi R/L)} \sin\left(\frac{\pi z}{L}\right) I_0\left(\frac{\pi r}{L}\right). \tag{63}$$

The boundary condition set along the line $r = R$ for potential (63) is a sinusoidal function with the amplitude U_R (Fig. 9, *a*). However, the potential along the horizontal line $r = R$ for real geometry with infinitely thin circular apertures must be exactly equal to $+U_R$ at the points

$$z_{2k} = (L/2) + 2Lk,$$

must be exactly equal to $-U_R$ at the points

$$z_{2k+1} = (3L/2) + 2Lk,$$

and must be a smooth monotonic function at intermediate points $z_{2k} \le z \le z_{2k+1}$, antisymmetric with respect to the central point located between the reference points z_{2k} and z_{2k+1}. In particular, it is acceptable to use a piecewise linear function as a model edge distribution of the potential (Fig. 9,*b*).

The boundary condition for the electric potential, given along the horizontal line $r = R$, is obviously a periodic function of the z coordinate with a period of $2L$ and therefore can be expanded into a Fourier series. Due to the symmetry of the geometric configuration of the electrodes and the antisymmetry of the potentials applied to the electrodes, only odd sinusoidal harmonics will be present in the Fourier series expansion, whereas cosine harmonics are strictly equal to zero.

For example, the considered Fourier series for the piecewise linear potential distribution shown in Fig. 9,*b* takes the form

$$U_\rho(z) \approx \frac{8U_R}{\pi^2}\sin\left(\frac{\pi z}{L}\right) - \frac{8U_R}{9\pi^2}\sin\left(\frac{3\pi z}{L}\right) + \frac{8U_R}{25\pi^2}\sin\left(\frac{5\pi z}{L}\right) - \frac{8U_R}{49\pi^2}\sin\left(\frac{7\pi z}{L}\right) + \ldots \tag{64}$$

The exact analytical solution for boundary condition (64) has the form:

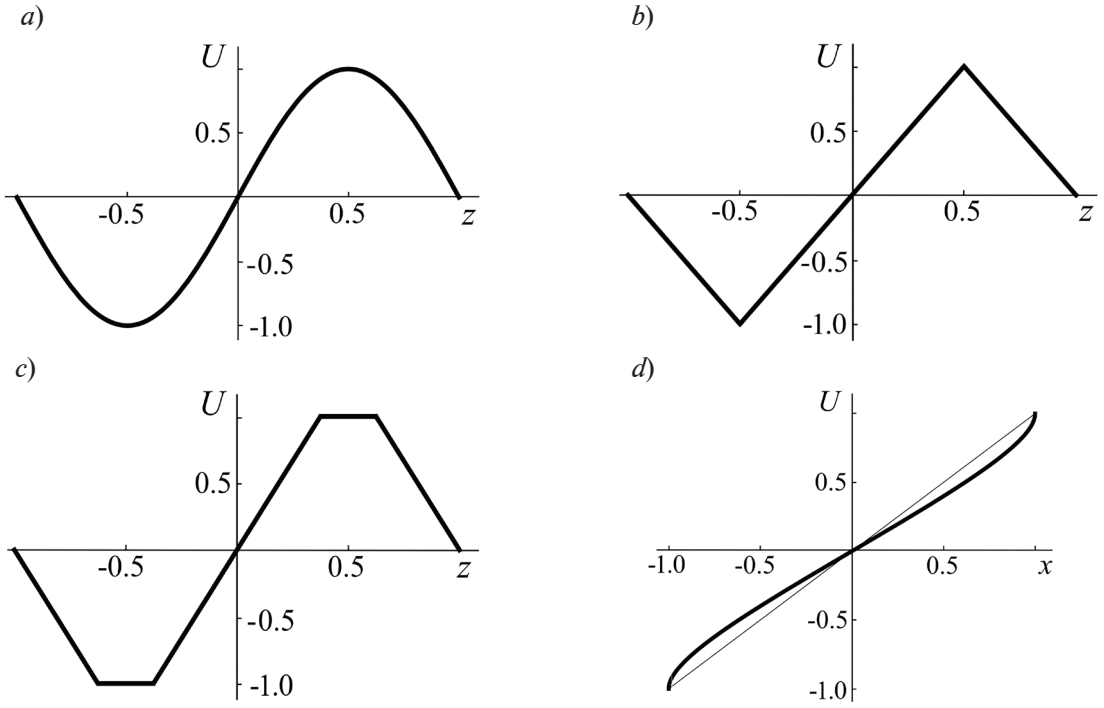

Fig. 9. Boundary conditions (in normalized coordinates) along the line $r = R$ in the interval between two adjacent apertures for calculation of electric field in SRIG-type trap for various model cases: *a* corresponds to analytical solution (9); *b* to infinitely thin apertures; *c* to apertures of finite thickness; *d* to substituting the linear approximation of the boundary value of the potential with an accurate analytical solution (see example 5 in paragraph 49, § 2, Chapter III in book [34])

$$U_S^*(z,r) = \frac{8U_R}{\pi^2 I_0(\pi R/L)} \sin\left(\frac{\pi z}{L}\right) I_0\left(\frac{\pi r}{L}\right) - \frac{8U_R}{9\pi^2 I_0(3\pi R/L)} \sin\left(\frac{3\pi z}{L}\right) I_0\left(\frac{3\pi r}{L}\right) + \\ + \frac{8U_R}{25\pi^2 I_0(5\pi R/L)} \sin\left(\frac{5\pi z}{L}\right) I_0\left(\frac{5\pi r}{L}\right) - \frac{8U_R}{49\pi^2 I_0(7\pi R/L)} \sin\left(\frac{7\pi z}{L}\right) I_0\left(\frac{7\pi r}{L}\right) + \cdots \quad (65)$$

The following equation holds true for the ratio of the amplitude of the third (parasitic) spatial harmonic to the amplitude of the first (main) spatial harmonic at a distance r from the symmetry axis in accordance with expression (65):

$$\Lambda(r) = \frac{1}{9} \frac{I_0(3\pi r/L)}{I_0(3\pi R/L)} \Big/ \frac{I_0(\pi r/L)}{I_0(\pi R/L)}. \quad (66)$$

Fig. 10 shows the dependences of Λ on the dimensionless quantity r/L for different values of the dimensionless geometric parameter R/L. It follows from these graphs, for example, that amplitude of the third harmonic for $R/L \geq 2$ does not exceed 5% of the amplitude of the first harmonic if $r/L \leq 1.75$. (The thin horizontal line in Fig. 10 corresponds to the 5% level for the quantity Λ). The amplitudes of the remaining parasitic spatial harmonics will be significantly smaller, and their influence can be ignored.

Notably, the relationship between the oscillation amplitude U_0 of the static electric potential on the axis of the system and the voltages U_R applied to the circular apertures for the «ideal» solution (63) is established by

$$U_0 = U_R / I_0(\pi R/L).$$

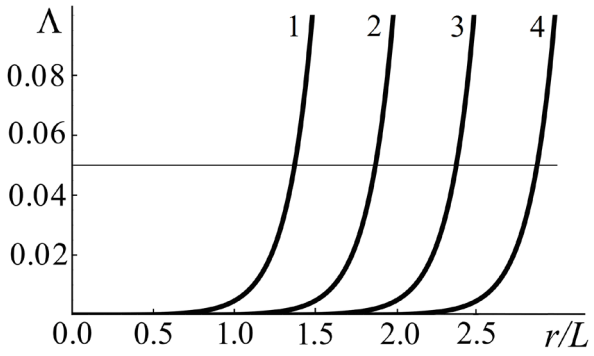

Fig. 10. Dependence of quantity Λ (see Eq. (66)) on dimensionless quantity r/L for different values of geometric parameter R/L: $R/L = 1.5$ (curve *1*), $R/L = 2.0$ (*2*), $R/L = 2.5$ (*3*), $R/L = 3.0$ (*4*). The horizontal line shows the level of 0.05 (5%)

On the other hand, this relationship for model Fourier series (64) is established as

$$U_0 = \left(8/\pi^2\right) U_R / I_0\left(\pi R/L\right).$$

Note. A piecewise linear function used in the intervals between reference points does not fully yield the required accuracy, however, it is rather acceptable for obtaining an approximate estimate of the sufficient distance to the edges of the electrodes. If it is necessary to take into account a more realistic model distribution of the potential along the horizontal line $r = R$, an analytical solution can be used for the boundary field of a parallel-plate capacitor, given in book [34] (see example 5 in paragraph 49, § 2, Chapter III of this monograph). This analytically accurate potential distribution along the cross section $y = 0$ at the edge of a parallel-plate capacitor is shown in Fig. 9,*d*; it follows that the piecewise linear function is largely acceptable as a model for estimating the contribution of the initial correction terms of the Fourier series along the horizontal line $r = R$. In particular, if the circular apertures are «thick» (Fig. 9,*c*), then the graphs in Fig. 9,*b* become piecewise trapezoidal within the framework of the considered piecewise linear model, and Fourier series (64) takes the following form:

$$U_\rho(z) \approx \frac{8\cos\left(\pi\gamma/2\right)U_R}{\pi^2\left(1-\gamma\right)}\sin\left(\frac{\pi z}{L}\right) - \frac{8\cos\left(3\pi\gamma/2\right)U_R}{9\pi^2\left(1-\gamma\right)}\sin\left(\frac{3\pi z}{L}\right) + \\ + \frac{8\cos\left(5\pi\gamma/2\right)U_R}{25\pi^2\left(1-\gamma\right)}\sin\left(\frac{5\pi z}{L}\right) - \frac{8\cos\left(7\pi\gamma/2\right)U_R}{49\pi^2\left(1-\gamma\right)}\sin\left(\frac{7\pi z}{L}\right) + \cdots, \tag{67}$$

where γ is a dimensionless parameter calculated by the formula $\gamma = h/L \in [0,1]$ (h is the thickness of the aperture, L is the distance between the centers of neighboring apertures (see above)).

Interestingly, within the framework of this model, the amplitude of the first spatial harmonic at $\gamma \approx 1/4$ coincides with the amplitude of the "ideal" case (63). However, the optimal choice here is the case $\gamma = 1/3$ when the third parasitic spatial harmonic vanishes.

Similarly, the rigorous shape of thin apertures corresponding to analytical solutions (13)–(16), (37)–(40) and (51)–(62) for segmented multipole transport channels differs from identical circular arcs with insulating gaps between them shown in Figs. 3, 4. A similar approach can be used to estimate how fast parasitic higher-order harmonics decay with distance away from the edges of segmented multipole circular apertures, so analytical formulas (13)–(16), (37)–(40) and (51)–(62) will be sufficient to describe the electric potential of the corresponding electric field with high accuracy.

As an example, consider a circular segmented quadrupole trap, for the electric field of which it was previously proposed to use analytical expressions (7)–(10), equivalent to each other up to rotation relative to the axis and displacement along the axis.

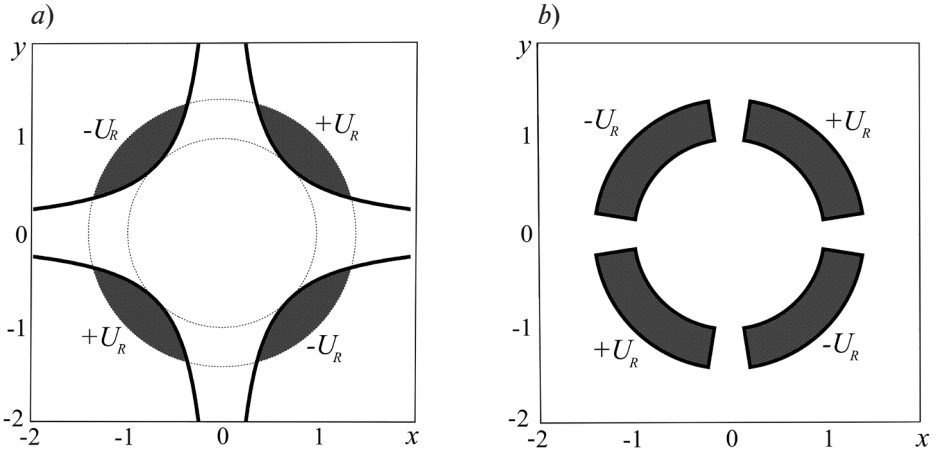

Fig. 11. Shape of individual elements of thin aperture for quadrupole segmentation of circular electrodes: *a* corresponds to analytical solution (9); *b* to circular segmented aperture

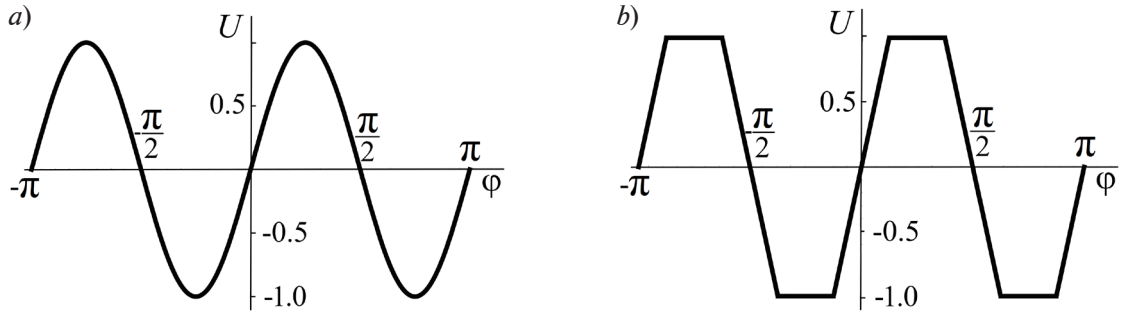

Fig. 12. Model angular distributions of potential in apertures under two boundary conditions: for analytical solution (9) (*a*) and for circular segmented quadrupole apertures (*b*)
The boundary conditions are set along the circle $r = R$ in the plane of the thin circular aperture of the SRIG-type RF trap with quadrupole segmentation of electrodes

Fig. 11,*a* shows the shape of the electrodes of thin apertures corresponding to analytical expression (9), Fig. 11,*b* shows those corresponding to thin circular segmented quadrupole apertures.

Consider circles $x = R\cos\varphi$ and $y = R\sin\varphi$ with radii $r = R$ in cross-sections $z = z_k$ with thin apertures to which electrical voltages alternating in sign ($z_{2k} = 2kL$ and $z_{2k+1} = (2k+1)L$) are applied.

Potential (9) is expressed along these circles by a sinusoidal function of the angular coordinate (Fig. 12,*a*). Fig. 12,*b* shows the model potential distribution for the circular segmented quadrupole aperture of radius R, where an approximate linear function is used instead of the true sinusoidal potential for the cross-sections of the circle corresponding to the insulating gaps. Expansion of this boundary condition into a Fourier series with respect to the angular coordinate

$$\varphi = \mathrm{Arg}(x, y) \in [-\pi, +\pi]$$

produces the expression

$$U_f(\varphi) \approx \frac{8\cos(\pi\delta/2)U_R}{\pi^2(1-\delta)}\sin(2\varphi) - \frac{8\cos(3\pi\delta/2)U_R}{9\pi^2(1-\delta)}\sin(6\varphi) + \\ + \frac{8\cos(5\pi\delta/2)U_R}{25\pi^2(1-\delta)}\sin(10\varphi) - \frac{8\cos(7\pi\delta/2)U_R}{49\pi^2(1-\delta)}\sin(14\varphi) + \cdots, \tag{68}$$

where δ is the relative angular size of a single multipole segment, calculated by the formula $\delta = 2\beta/\pi \in [0,1]$ ($\beta \approx (\pi/2) - (\Delta/R)$ is the absolute angular size of the given segment, Δ is the

gap between the segments, R is the radius of the circular aperture, $\pi/2$ is the angular distance between the centers of circular multipole segments in the case of quadrupole segmentation).

The coefficients of Fourier series (68) coincide with the coefficients of Fourier series (67). In view of this, despite the different physical meanings of the quantities δ and γ, if $\delta \approx 1/4$, the amplitude of the first angular spatial harmonic coincides with the amplitude of the "ideal" case (9). However, as in the previous case, the optimal choice is the value $\delta = 1/3$ when the third parasitic angular harmonic vanishes.

The exact analytical solution corresponding to boundary condition (68) takes the form:

$$U_C^{(q,*)}(z,x,y) = \frac{8U_R}{\pi^2(1-\delta)}\left[\frac{\cos(\pi\delta/2)}{I_2(\pi R/L)}\frac{Q_2(x,y)}{r^2}I_2\left(\frac{\pi\sqrt{x^2+y^2}}{L}\right)-\right.$$
$$-\frac{\cos(3\pi\delta/2)}{9I_6(\pi R/L)}\frac{Q_6(x,y)}{r^6}I_6\left(\frac{\pi\sqrt{x^2+y^2}}{L}\right)+$$
$$+\frac{\cos(5\pi\delta/2)}{25I_{10}(\pi R/L)}\frac{Q_{10}(x,y)}{r^{10}}I_{10}\left(\frac{\pi\sqrt{x^2+y^2}}{L}\right)-$$
$$\left.-\frac{\cos(7\pi\delta/2)}{49I_{14}(\pi R/L)}\frac{Q_{14}(x,y)}{r^{14}}I_{14}\left(\frac{\pi\sqrt{x^2+y^2}}{L}\right)+\cdots\right]\cos\left(\frac{\pi z}{L}\right), \quad (69)$$

where the multipole multipliers $Q_k(x,y)$ are expressed as follows in accordance with recurrence relations (25)–(27):

$$Q_2(x,y) = 2xy,$$
$$Q_6(x,y) = 2xy\left(x^2-3y^2\right)\left(3x^2-y^2\right),$$
$$Q_{10}(x,y) = 2xy\left(x^4-10x^2y^2+5y^4\right)\left(5x^4-10x^2y^2+y^4\right),$$
$$Q_{14}(x,y) = 2xy\left(x^6-21x^4y^2+35x^2y^4-7y^6\right)\left(7x^6-35x^4y^2+21x^2y^4-y^6\right).$$

The ratio of the amplitude of the sixth (parasitic) spatial angular harmonic to the amplitude of the second (main) spatial angular harmonic is expressed as follows at a distance r from the symmetry axis, in accordance with expression (69),

$$\Omega(r) = \frac{1}{9}\frac{\cos(3\pi\delta/2)}{I_6(\pi R/L)}I_6\left(\frac{\pi r}{L}\right)\Big/\frac{\cos(\pi\delta/2)}{I_2(\pi R/L)}I_2\left(\frac{\pi r}{L}\right). \quad (70)$$

Fig. 13 shows graphs for the dependence of Ω on the dimensionless ratio r/L for different values of the geometric parameter R/L; the value $\delta = 1/5$ is used in this example[1]. In particular, it follows from the graphs that the amplitude of the sixth spatial angular harmonic for $R/L \geq 2$ does not exceed 5% of the amplitude of the second spatial angular harmonic if $r/L \leq 1.75$. (As in Fig. 10, the thin horizontal line in Fig. 13 corresponds to the 5% level of the quantity Ω.) The amplitudes of the remaining parasitic spatial angular harmonics will be significantly smaller and, as in the previous case, their influence can be neglected.

Notably, the relationship between the oscillation amplitude U_0 of the static electric potential on the axis of the system and the quadrupole voltages $\pm U_R^{(r)}$ applied to the quadrupole segments of thin circular apertures for the "ideal" solution (9) is established by the equality

[1] Varying the parameter $\delta \in [0,1]$ allows to change the factor multiplying Ω from 0 at $\delta = 1/3$ to 3 at $\delta = 1$. This multiplier is 0.618 for $\delta = 1/5$. Naturally, at $\delta = 1/3$, when the sixth harmonic vanishes, it is necessary to evaluate the ratio of the tenth harmonic to the second harmonic.

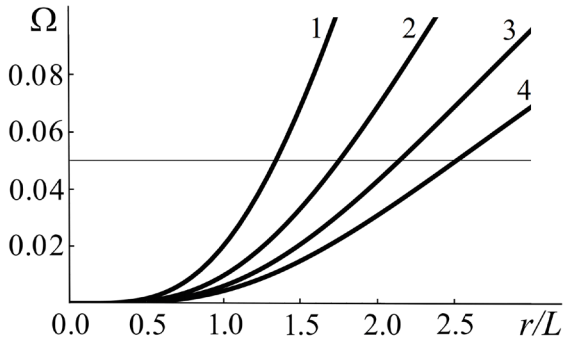

Fig. 13. Dependences of Ω (see Eq. (70)) on dimensionless ratio r/L for different values of geometric parameter R/L:
R/L = 1.5 (curve *1*), R/L = 2.0 (*2*), R/L = 2.5 (*3*), R/L = 3.0 (*4*).
The horizontal line shows the level of 0.05 (5%)

$$U_0 = U_R^{(r)} \big/ I_2\left(\pi R/L\right).$$

At the same time, this relationship for model Fourier series (68) is established by the relation

$$U_0 = \left(8\cos\left(\pi\gamma/2\right)\big/\pi^2\left(1-\gamma\right)\right)U_R^{(r)} \big/ I_2\left(\pi R/L\right).$$

The distance from the edges of the electrodes where it is acceptable to use the analytical expressions given in this paper for conical RF funnels and RF funnels with curved profiles is estimated by a similar procedure.

Importantly, the boundary condition for analysis of conical RF funnels and RF funnels with curved profiles should be imposed using the envelope of the inner edge of the funnel rather than the straight line $r = R$. This envelope is determined from the condition that the analytical expression at the points $\left(z_k, r_k\right)$ is equal to the electric voltage applied to the aperture and is the same (up to a sign) for all apertures.

Conclusion

Analytical expressions for electrical potentials obtained in this paper can be used for analyzing ion motion in radio frequency traps and radio frequency funnels incorporating circular apertures (in particular, segmented multipole apertures) with polynomial profiles. For instance, analytical expressions can be used to quickly qualitatively investigate and optimize the behavior of ions in these devices using a pseudopotential model of ion motion in high-frequency electric fields [35–37].

Considering weighted sums of analytical expressions that correspond to circular apertures and segmented multipole apertures placed along the axis with the same pitch allows to model the behavior of ions in traps and funnels with non-segmented sections different from the circular shape. If these sums correspond to the cases of apertures placed along the axis with multiple pitch and/or multiple multipole segmentation of electrodes, then it is possible to investigate the influence of parasitic spatial angular harmonics of the electric field induced by the imperfect geometric shape of the electrodes on ion motion.

The obtained analytical expressions for three-dimensional harmonic functions with polynomial oscillating behavior on the axis can also be useful in solving certain problems of mathematical physics.

Acknowledgment

The authors express their sincere gratitude to Mikhail Igorevich Yavor, Dr. Sci. (phys.-math.), Principal Investigator at the Institute for Analytical Instrumentation of the Russian Academy of Sciences (St. Petersburg), for active participation in the discussion of the problem under consideration and pointing to useful literary sources for the study.

Wolfram Mathematica version 11 (Home Edition) [38] was used for the calculations.

REFERENCES

1. **Sysoyev A. A., Berdnikov A. S., Masyukevich S. V., et al.,** Analytical study of operating modes of RF ion funnels in the gas dynamic interfaces of tandem triple-quadrupole mass-spectrometers, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 6 (4) (2023) 134–145 (in Russian).

2. **Gerlich D.,** Inhomogeneous RF fields: A versatile tool for the study of processes with slow ions, In book: Ng Ch.-Y., Baer M. (Eds.). State-selected and state-to-state ion–molecule reaction dynamics. Part 1: Experiment (Book Series: Advances in Chemical Physics. Vol. LXXXII). John Wiley & Sons Inc., New York (1992) 1–176.

3. **Yavor M. I.,** Optics of charged particle analyzers (Book Series: Advances in Imaging and Electron Physics. Vol. 157), Academic Press, Amsterdam (2009) 142–168.

4. **Bahr R.** Diplom Thesis. University of Freiburg. 1969 (cited by Refs. [2, 6, 7]).

5. **Gerlich D.** Diplom Thesis. University of Freiburg. 1971 (cited by Refs. [2, 6, 7]).

6. **Teloy E., Gerlich D.,** Integral cross sections for ion-molecule reactions. Part I. The Guided beam technique, Chem. Phys. 4 (3) (1974) 417–427.

7. **Gerlich D., Kaefer G.,** Ion trap studies of association processes in collisions of CH^{3+} and CD^{3+} with *n*-H_2, *p*-H_2, D_2 and He at 80 K, Astrophys. J. 347 (2) (1989) 849–854.

8. **Berdnikov A. S., Gall N. R.,** Radio frequency ion guiding traps with periodic electrodes without spurious trapping regions, J. Anal. Chem. 69 (13) (2014) 1285–1290.

9. **Verentchikov A., Berdnikov A., Yavor M.,** Stable ion beam transport through periodic electrostatic structures and non-linear effects, Phys. Procedia 1 (1) (2008) 87–97.

10. **Shaffer S. A., Tang K., Anderson G., et al.,** A novel ion funnel for focusing ions at elevated pressure using electrospray ionization mass spectrometry, Rapid Commun. Mass Spectrom. 11 (16) (1997) 1813–1817.

11. **Shaffer S. A., Prior D. C., Anderson G., et al.,** An ion funnel interface for improved ion focusing and sensitivity using electrospray ionization mass spectrometry, Anal. Chem. 70 (19) (1998) 4111–4119.

12. **Shaffer S. A., Tolmachev A. V., Prior D. C., et al.,** Characterization of an improved electrodynamic ion funnel interface for electrospray ionization mass spectrometry, Anal. Chem. 71 (15) (1999) 2957–2964.

13. **Tolmachev A. V., Kim T., Udseth H. R., et al.,** Simulation-based optimization of the electrodynamic ion funnel for high sensitivity electrospray ionization mass spectrometry, Int. J. Mass Spectrom. 203 (1–3) (2000) 31–47.

14. **Kim T., Tolmachev A. V., Harkewicz R., et al.,** Design and implementation of a new electrodynamic ion funnel, Anal. Chem. 72 (10) (2000) 2247–2255.

15. **Lynn E. C., Chung M.-C., Han C.-C.,** Characterizing the transmission properties of an ion funnel, Rapid Commun. Mass Spectrom. 14 (22) (2000) 2129–2134.

16. **Kelly R. T., Tolmachev A. V., Page J. S., et al.,** The ion funnel: Theory, implementations and applications, Mass Spectrom. Rev. 29 (2) (2010) 294–312.

17. **Watson G. N.,** A treatise on the theory of Bessel functions, 2nd Ed., Cambridge University Press, Cambridge, UK, 1945.

18. **Bateman H., Erdélyi A.,** Higher transcendental functions, Vol. II, McGraw-Hill Book Company, New York, Toronto, London, 1953.

19. **Abramowitz M., Stegun I. A.** (Eds.), Handbook on mathematical functions with formulas, graphs, and mathematical tables, 10th edition (Applied Mathematics Series, Vol. 55), USA National Bureau of Standards, 1972.

20. **Andreyeva A. D., Berdnikov A. S.,** Ustroystva dlya manipulirovaniya zaryazhennymi chastitsami na osnove printsipa arkhimedova vinta [Devices for manipulating charged particles based on the principle of Archimedean screw], Proc. of the 4-th All-Russian conf. "Mass-Spectrometry and its Applied Problems" organized by the 5-th Congress of All-Russian Mass-Spectrometry Society; Sept. 05–09, 2011, Moscow (2011) 137 (in Russian).

21. **Andreyeva A. D., Berdnikov A. S.,** Mass spectrometric devices with Archimedean radio-frequency electric fields, J. Anal. Chem. 67 (13) (2012) 1034–1037.

22. **Berdnikov A. S., Douglas D. J., Konenkov N. V.,** The pseudopotential for quadrupole fields up to $q = 0.9080$, Int. J. Mass Spectrom. 421 (2017) 204–223.

23. **Berdnikov A. S., Andreeeva A. D.,** Ustroystvo dlya manipulirovaniya zaryazhennymi chastitsami, Patent Federalnoy sluzhby RF po intellektualnoy sobstvennosti na poleznuyu model RU 113611 (Data prioriteta/podchi 05.05.2011) [A device for charge particle manipulation. Patent of the Federal Service for Intellectual Property for a utility model of the Russian Federation] – RU 113611 (Priority/ Application date: 05.05.2011) (in Russian).

24. **Berdnikov A. S., Andreeeva A. D.,** Ustroystvo dlya manipulirovaniya zaryazhennymi chastitsami, Patent Federalnoy sluzhby RF po intellektualnoy sobstvennosti na izobreteniye RU 113611 (Data prioriteta/podchi 05.05.2011) [A device for charge particle manipulation. Patent of the Federal Service for Intellectual Property for an invention of the Russian Federation] – RU 2465679 (Priority/ Application date: 05.05.2012) (in Russian).

25. **Berdnikov A., Andreyeva A., Giles R.,** Device for manipulating charged particles via field with pseudo potential having one or more local maxima along length of channel, Patent US9536721 (2017).

26. **Berdnikov A., Andreyeva A., Giles R.**, Device for manipulating charged particles, Patent US9812308 (2017).

27. **Berdnikov A. S.,** Time-dependent pseudopotential and its application for description of the averaged motion of the charged particles. Part. I, Nauchnoe priborostroenie [Scientific Instrumentation]. 21 (2) (2011) 77–89 (in Russian).

28. **Berdnikov A. S.,** Time-dependent pseudopotential and its application for description of the averaged motion of the charged particles. Part 2. General expression for time-dependent pseudopotentials, Nauchnoe priborostroenie [Scientific Instrumentation]. 21 (3) (2011) 83–96 (in Russian).

29. **Berdnikov A. S.,** Time-dependent pseudopotential and its application for description of the averaged motion of the charged particles. Part 3. Time dependent signals characterized by "slow" and "fast" characteristic time, Nauchnoe priborostroenie [Scientific Instrumentation]. 21 (4) (2011) 75–85 (in Russian).

30. **Berdnikov A. S.,** Time-dependent pseudopotential and its application for description of the averaged motion of the charged particles. Part 4. Devices and instruments, Nauchnoe priborostroenie [Scientific Instrumentation]. 21 (4) (2011) 86–102 (in Russian).

31. **Berdnikov A. S.,** Time-dependent pseudopotential and its application for description of the averaged motion of the charged particles. Part 5. Comments to a general expression for time-dependent pseudopotentials, Nauchnoe priborostroenie [Scientific Instrumentation]. 22 (2) (2012) 105–111 (in Russian).

32. **Berdnikov A. S.,** High frequency electromagnetic fields with Archimedean properties, Nauchnoe priborostroenie [Scientific Instrumentation]. 24 (1) (2014) 104–127 (in Russian).

33. **Bao X., Zhang Q., Liang Q., et al.,** Increased sensitivity in proton transfer reaction mass spectrometry by using a novel focusing quadrupole ion funnel, Anal. Chem. 94 (39) (2022) 13368–13376.

34. **Lavrentiev M., Shabat B.,** Methodes de la Theorie des Fonctions d'une Variable Complexe, "Mir" D'Édition, Moscou, 1972.

35. **Sagdeev R. Z., Usikov D. A., Zaslavski G. M.,** Nonlinear physics: From the pendulum to turbulence and chaos (Contemporary Concepts in Physics, Vol. 4), Harwood Academic Publishers, Chur, Switzerland; N. Y., 1988.

36. **Berdnikov A. S.,** A pseudo potential description of the motion of charged particles in RF fields, Microsc. Microanal. 21 (S4) (2015) 78–83.

37. **Berdnikov A. S., Kuzmin A. G., Verenchikov A. N.**, On the correct averaging of the equations of ion motion in high-frequency electric fields, J. Anal. Chem. 74 (14) (2019) 1378–1389.

38. Wolfram Mathematica: The system for modern technical computing, URL: http://wolfram.com/mathematica/

СПИСОК ЛИТЕРАТУРЫ

1. **Сысоев А. А., Бердников А. С., Масюкевич С. В., Соловьев К. В., Краснова Н. К.** Аналитическое исследование режимов работы радиочастотных воронок в газодинамических интерфейсах тандемных трехквадрупольных масс-спектрометров // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 16. № 4. С. 134–145.

2. **Gerlich D.** Inhomogeneous RF fields: A versatile tool for the study of processes with slow ions // Ng Ch.-Y., Baer M. (Eds.). State-selected and state-to-state ion–molecule reaction dynamics. Part 1: Experiment (Book Series: Advances in Chemical Physics. Vol. LXXXII). New York: John Wiley & Sons Inc., 1992. Pp. 1–176.

3. **Yavor M. I.** Optics of charged particle analyzers (Book Series: Advances in Imaging and Electron Physics. Vol. 157). Amsterdam: Academic Press, 2009. Pp. 142–168.

4. **Bahr R.** Diplom thesis. University of Freiburg. 1969 (cited by Refs. [2, 6, 7]).

5. **Gerlich D.** Diplom thesis. University of Freiburg. 1971 (cited by Refs. [2, 6, 7]).

6. **Teloy E., Gerlich D.** Integral cross sections for ion-molecule reactions. Part I. The Guided beam technique // Chemical Physics. 1974. Vol. 4. No. 3. Pp. 417–427.

7. **Gerlich D., Kaefer G.** Ion trap studies of association processes in collisions of CH^{3+} and CD^{3+} with *n*-H_2, *p*-H_2, D_2 and He at 80 K // The Astrophysical Journal. 1989. Vol. 347. No. 2. Pp. 849–854.

8. **Бердников А. С., Галль Н. Р.** Радиочастотные транспортирующие ловушки с периодическими электродами без паразитных областей захвата // Масс-спектрометрия. 2013. Т. 10. № 4. С. 224–229.

9. **Verentchikov A., Berdnikov A., Yavor M.** Stable ion beam transport through periodic electrostatic structures and non-linear effects // Physics Procedia. 2008. Vol. 1. No. 1. Pp. 87–97.

10. **Shaffer S. A., Tang K., Anderson G., Prior D. C., Udseth H. R., Smith R. D.** A novel ion funnel for focusing ions at elevated pressure using electrospray ionization mass spectrometry // Rapid Communications in Mass Spectrometry. 1997. Vol. 11. No. 16. Pp. 1813–1817.

11. **Shaffer S. A., Prior D. C., Anderson G., Udseth H. R., Smith R. D.** An ion funnel interface for improved ion focusing and sensitivity using electrospray ionization mass spectrometry // Analytical Chemistry. 1998. Vol. 70. No. 19. Pp. 4111–4119.

12. **Shaffer S. A., Tolmachev A. V., Prior D. C., Anderson G. A., Udseth H. R., Smith R. D.** Characterization of an improved electrodynamic ion funnel interface for electrospray ionization mass spectrometry // Analytical Chemistry. 1999. Vol. 71. No. 15. Pp. 2957–2964.

13. **Tolmachev A. V., Kim T., Udseth H. R., Smith R. D., Bailey T. H., Futrell J. H.** Simulation-based optimization of the electrodynamic ion funnel for high sensitivity electrospray ionization mass spectrometry // International Journal of Mass Spectrometry. 2000. Vol. 203. No. 1–3. Pp. 31–47.

14. **Kim T., Tolmachev A. V., Harkewicz R., Prior D. C., Anderson G., Udseth H. R., Smith R. D.** Design and implementation of a new electrodynamic ion funnel // Analytical Chemistry. 2000. Vol. 72. No. 10. Pp. 2247–2255.

15. **Lynn E. C., Chung M.-C., Han C.-C.** Characterizing the transmission properties of an ion funnel // Rapid Communications in Mass Spectrometry. 2000. Vol. 14. No. 22. Pp. 2129–2134.

16. **Kelly R. T., Tolmachev A. V., Page J. S., Tang K., Smith R. D.** The ion funnel: Theory, implementations and applications // Mass Spectrometry Reviews. 2010. Vol. 29. No. 2. Pp. 294–312.

17. **Ватсон Дж. Н.** Теория бесселевых функций. Пер. со -2го англ. изд. М.: Издательство иностранной литературы, 799 .1949 с.

18. **Бейтмен Г., Эрдейи А.** Высшие трансцендентные функции. Т. 2: Функции Бесселя, функции параболического цилиндра, ортогональные многочлены. Серия «Справочная математическая библиотека». Изд. -2е. М.: Наука, 296 .1974 с.

19. **Абрамовиц М., Стиган И.** Справочник по специальным функциям с формулами, графиками и математическими таблицами. М.: Наука, 832 .1979 с.

20. **Андреева А. Д., Бердников А. С.** Устройства для манипулирования заряженными частицами на основе принципа архимедова винта // Сборник тезисов докладов IV Всероссийской конференции и V съезда Всероссийского масс-спектрометрического общества «Масс-спектрометрия и ее прикладные проблемы» (Москва, 09–05 сентября 2011 г.). 2011. С. 137.

21. **Андреева А. Д., Бердников А. С.** Масс-спектрометрические устройства на основе радиочастотных электрических полей с архимедовыми свойствами // Масс-спектрометрия. 2011. Т. 8. № 4. С. 293–296.

22. **Berdnikov A. S., Douglas D. J., Konenkov N. V.** The pseudopotential for quadrupole fields up to $q = 0.9080$ // International Journal of Mass Spectrometry. 2017. Vol. 421. Pp. 204–223.

23. **Бердников А. С., Андреева А. Д.** Устройство для манипулирования заряженными частицами. 2011. Патент Федеральной службы Российской Федерации по интеллектуальной собственности на полезную модель RU 113611 (Дата приоритета/подачи заявки: 05.05.2011, дата публикации: 20.02.2012).

24. **Бердников А. С., Андреева А. Д.** Устройство для манипулирования заряженными частицами. 2012. Патент Федеральной службы Российской Федерации по интеллектуальной собственности на изобретение RU 2465679 (Дата приоритета/подачи заявки: 05.05.2011, дата публикации: 27.10.2012).

25. **Berdnikov A., Andreyeva A., Giles R.** Device for manipulating charged particles via field with pseudo potential having one or more local maxima along length of channel // Patent US9536721 (Priority date: 05.05.2011, Application date: 04.05.2012, Publication date: 03.01.2017).

26. **Berdnikov A., Andreyeva A., Giles R.** Device for manipulating charged particles // Patent US9812308. (Priority date: 05.05.2011, Application date: 21.10.2016, Publication date: 07.11.2017).

27. **Бердников А. С.** Меняющийся во времени псевдопотенциал и его применение к описанию усредненного движения заряженных частиц. Ч. I. // Научное приборостроение. 2011. Т. 21. № 2. С. 77–89.

28. **Бердников А. С.** Меняющийся во времени псевдопотенциал и его применение к описанию усредненного движения заряженных частиц. Ч. 2. Общая формула // Научное приборостроение. 2011. Т. 3 № .21. С. 96–83.

29. **Бердников А. С.** Меняющийся во времени псевдопотенциал и его применение к описанию усредненного движения заряженных частиц. Ч. 3. Временные сигналы, характеризуемые «медленным» и «быстрым» временами // Научное приборостроение. 2011. Т. 4 № .21. С. 85–75.

30. **Бердников А. С.** Меняющийся во времени псевдопотенциал и его применение к описанию усредненного движения заряженных частиц. Ч. 4. Приборы и устройства // Научное приборостроение. 2011. Т. 4 № .21. С. 102–86.

31. **Бердников А. С.** Меняющийся во времени псевдопотенциал и его применение к описанию усредненного движения заряженных частиц. Ч. 5. Комментарии к общей формуле для меняющихся во времени псевдопотенциалов // Научное приборостроение. 2012. Т. 22. № 2. С. 105–111.

32. **Бердников А. С.** Высокочастотные электромагнитные поля с архимедовыми свойствами // Научное приборостроение. 2014. Т. 24. № 1. С. 104–127.

33. **Bao X., Zhang Q., Liang Q., et al.** Increased sensitivity in proton transfer reaction mass spectrometry by using a novel focusing quadrupole ion funnel // Analytical Chemistry. 2022. Vol. 94. No. 39. Pp.13368–13376.

34. **Лаврентьев М. А., Шабат Б. В.** Методы теории функций комплексного переменного. Изд. -3е, испр. М.: «Наука». Гл. ред. физико-математической лит-ры, 1965. 716 с.

35. **Заславский Г. М., Сагдеев Р. З.** Введение в нелинейную физику. От маятника до турбулентности и хаоса. М.: “Наука». Гл. ред. физико-математической лит-ры, 1988. 368 с.

36. **Berdnikov A. S.** A pseudo potential description of the motion of charged particles in RF fields // Microscopy and Microanalysis. 2015. Vol. 21. No. S4. Pp. 78–83.

37. **Бердников А. С., Веренчиков А. Н., Кузьмин А. Г.** О корректном усреднении уравнений движения ионов в высокочастотных электрических полях // Масс-спектрометрия. 2018. Т. 15. № 4. С. 233–245.

38. Wolfram Mathematica: the system for modern technical computing: URL: http://wolfram.com/mathematica/

THE AUTHORS

BERDNIKOV Alexander S.
Institute for Analytical Instrumentation, RAS
31–33, Ivan Chernykh St., St. Petersburg, 198095, Russia
asberd@yandex.ru
ORCID: 0000-0003-0985-5964

MASYUKEVICH Sergey V.
Institute for Analytical Instrumentation, RAS
31–33, Ivan Chernykh St., St. Petersburg, 198095, Russia
serg_08@mail.ru
ORCID: 0000-0002-0873-8849

POMOZOV Timofey V.
Institute for Analytical Instrumentation, RAS
31–33, Ivan Chernykh St., St. Petersburg, 198095, Russia
tim-kor@mail.ru
ORCID: 0000-0001-8925-7600

HASIN Yuri I.
Institute for Analytical Instrumentation, RAS
31–33, Ivan Chernykh St., St. Petersburg, 198095, Russia
yur-khasin@yandex.ru
ORCID: 0009-0002-1506-0631

SOLOVYEV Konstantin V.
Peter the Great St. Petersburg Polytechnic University
Institute for Analytical Instrumentation, RAS
29, Politechnicheskaya St., St. Petersburg, 195251, Russia
k-solovyev@mail.ru
ORCID: 0000-0003-3514-8577

СВЕДЕНИЯ ОБ АВТОРАХ

БЕРДНИКОВ Александр Сергеевич – *доктор физико-математических наук, главный научный сотрудник Института аналитического приборостроения РАН.*
198095, Россия, г. Санкт-Петербург, ул. Ивана Черных, 31–33, лит. А.
asberd@yandex.ru
ORCID: 0000-0003-0985-5964

МАСЮКЕВИЧ Сергей Владимирович – *старший научный сотрудник Института аналитического приборостроения РАН.*
198095, Россия, г. Санкт-Петербург, ул. Ивана Черных, 31–33, лит. А.
serg_08@mail.ru
ORCID: 0000-0002-0873-8849

ПОМОЗОВ Тимофей Вячеславович – *кандидат физико- математических наук, старший научный сотрудник Института аналитического приборостроения РАН.*
198095, Россия, г. Санкт-Петербург, ул. Ивана Черных, 31–33, лит. А.
tim-kor@mail.ru
ORCID: 0000-0001-8925-7600

ХАСИН Юрий Иосифович – *кандидат физико-математических наук, ведущий научный сотрудник Института аналитического приборостроения РАН.*
198095, Россия, г. Санкт-Петербург, ул. Ивана Черных, 31–33, лит. А.
yur-khasin@yandex.ru
ORCID: 0009-0002-1506-0631

СОЛОВЬЕВ Константин Вячеславович – *кандидат физико-математических наук, доцент Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого, младший научный сотрудник Института аналитического приборостроения Российской академии наук.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
k-solovyev@mail.ru
ORCID: 0000-0003-3514-8577

Received 20.01.2024. Approved after reviewing 26.02.2024. Accepted 26.02.2024.
Статья поступила в редакцию 20.01.2024. Одобрена после рецензирования 26.02.2024. Принята 26.02.2024.

PHYSICAL MATERIALS TECHNOLOGY

Original article
DOI: https://doi.org/10.18721/JPM.17209

THE FORMATION OF SINGLE-DOMAIN GALLIUM PHOSPHIDE BUFFER LAYERS ON A SILICON SUBSTRATE WITHOUT THE USE OF MIGRATION ENHANCED EPITAXY TECHNIQUE

V. V. Fedorov[1,2]✉, *S. V. Fedina*[1,2], *A. K. Kaveev*[3,2]

D. A. Kirilenko[3,4], *N. N. Faleev*[3], *I. S. Mukhin*[1,2]

[1] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia;

[2] Alferov University, RAS, St. Petersburg, Russia;

[3] Ioffe Institute, RAS, St. Petersburg, Russia;

[4] ITMO University, St. Petersburg, Russia

✉ fedorov_vv@spbstu.ru

Abstract. In the paper, the influence of growth factors and silicon surface condition on the epitaxial formation of single-domain GaP buffer layers on Si (001) has been studied. A novel two-stage growth technique for the epitaxial building-up of this structure has been put forward and developed. In contrast to using the migration enhanced epitaxy technique, the proposed technology allows one to separate the nucleation and growth stages, to control the doping profile of the GaP buffer layers. The latter is important for further functional applications. The main factors determining the orientation of GaP crystalline lattice when it nucleates on the Si vicinal surface were found. The structural perfection of the grown buffer layers at both stages was proved through careful control by TEM, RHEED, AFM, and HRXRD. These findings have important implications for further functional applications.

Keywords: heterostructure, molecular beam epitaxy, buffer layer, gallium phosphide, silicon substrate

Funding: The reported study was funded by Russian Science Foundation (Grant No. 22-29-01080), https://rscf.ru/project/22-29-01080.

Citation: Fedorov V. V., Fedina S. V., Kaveev A. K., Kirilenko D. A., Faleev N. N., Mukhin I. S., The formation of single-domain gallium phosphide buffer layers on a silicon substrate without the use of migration enhanced epitaxy technique, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 120–133. DOI: https://doi.org/10.18721/JPM.17209

Научная статья
УДК 539.23
DOI: https://doi.org/10.18721/JPM.17209

ФОРМИРОВАНИЕ ОДНОДОМЕННЫХ БУФЕРНЫХ СЛОЕВ ФОСФИДА ГАЛЛИЯ НА КРЕМНИЕВОЙ ПОДЛОЖКЕ БЕЗ ПРИМЕНЕНИЯ МЕТОДА ЭПИТАКСИИ С ПОВЫШЕННОЙ МИГРАЦИЕЙ

В. В. Федоров[1,2]✉, *С. В. Федина*[1,2], *А. К. Кавеев*[3,2]

Д. А. Кириленко[3,4], *Н. Н. Фалеев*[3], *И. С. Мухин*[1,2]

[1] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;

[2] Академический университет им. Ж. И. Алфёрова РАН, Санкт-Петербург, Россия;

[3] Физико-технический институт им. А. Ф. Иоффе РАН, Санкт-Петербург, Россия;

[4] Национальный исследовательский университет информационных технологий, механики и оптики (ИТМО), Санкт-Петербург, Россия

✉ fedorov_vv@spbstu.ru

Аннотация. В работе исследовано влияние ростовых условий и состояния поверхности кремния на процессы формирования буферных слоев фосфида галлия GaP на подложках кремния Si (001). Предложен и развит двухстадийный метод эпитаксиального выращивания псевдоморфных однодоменных буферных слоев GaP на Si (001), обеспечивающий разделение стадий зарождения и роста слоя на подложке. В отличие от метода эпитаксии с повышенной миграцией, предложенная технология позволяет управлять профилем легирования буферных слоев GaP на Si, что важно для дальнейших функциональных применений. Найдены основные факторы, определяющие ориентацию кристаллической решетки GaP при ее зарождении на вицинальной поверхности подложки. Путем тщательного контроля ростового процесса на обеих стадиях методами просвечивающей электронной микроскопии (TEM), дифракции быстрых электронов на отражение (RHEED), атомно-силовой микроскопии (AFM) и высокоразрешающей рентгеновской дифрактометрии (HRXRD) доказано высокое структурное совершенство выращенных буферных слоев.

Ключевые слова: гетероструктура, молекулярно-пучковая эпитаксия, буферный слой, фосфид галлия, кремниевая подложка

Финансирование: Исследование выполнено за счет гранта Российского научного фонда № 01080-29-22 (https://rscf.ru/project/22-29-01080).

Ссылка для цитирования: Федоров В. В., Федина С. В., Кавеев А. К., Кириленко Д. А., Фалеев Н. Н., Мухин И. С. Формирование однодоменных буферных слоев фосфида галлия на кремниевой подложке без применения метода эпитаксии с повышенной миграцией // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 2. С. 120–133. DOI: https://doi.org/10.18721/JPM.17209

Introduction

Integration of heterostructures based on $A^{III}B^{V}$ semiconductor compounds on silicon allows to combine the advantages of highly developed silicon-integrated technology with the unique capabilities of the technology for growing $A^{III}B^{V}$ heterostructures, such as the implementation of electronic and optical interconnects by constructing a lateral energy band profile of the structure [1].

A particular focus in the literature is on monolithic (epitaxial) integration of $A^{III}B^{V}$ heterostructures on silicon substrates. It is assumed that directly growing $A^{III}B^{V}$ makes it possible to expand the functionality of the resulting heterostructures and reduce the number of technological stages, compared with the purely hybrid technology (where preformed elements are transferred to silicon wafers).

However, high-precision epitaxial growth of $A^{III}B^{V}$ layers on silicon surfaces is associated with a number of obstacles characteristic for heteroepitaxy, including the following:

defects in the crystal structure, appearing due to inelastic relaxation of misfit stresses (the mismatch between the lattice parameters of the layer and the substrate);

symmetry mismatch between the crystal lattices of $A^{III}B^{V}$ layer (zinc-blende structure) and Si substrate (diamond structure), leading to the formation of antiphase domains and antiphase domain boundaries between them in the epitaxial layer;

undesirable surface chemical reactions at the heterointerface;

significant difference in surface energies of silicon, epitaxial layer and heterointerface.

One of the most significant systems comprises the epitaxial layers of gallium phosphide (GaP) on the surface of a Si (001) silicon substrate. Even though all binary $A^{III}B^{V}$ compounds exhibit the lowest mismatch of the GaP and Si crystal lattices ($\Delta a/a$ amounts to only about 0.37% for these structures), the crystalline perfection of GaAs layers recently grown on Ge(001) substrate ($\Delta a/a \approx 0.07\%$) or even GaAs layers grown through a metamorphic [2, 3] buffer on Si(001) ($\Delta a/a \approx 4\%$) often turns out to be higher than the quality of GaP layers on Si [4, 5]. The difficulties in growing gallium phosphide layers are due to the tendency of this material to the three-dimensional growth mechanism on the silicon surface and the possibility of undesirable surface chemical reactions. For example, phosphorus atoms can replace surface silicon atoms, disrupting the homogeneity of the surface [6, 7], while gallium can etch the silicon surface, leading to the formation of hollow voids [8].

Migration-enhanced epitaxy with alternate exposure of the substrate to phosphorus and gallium fluxes is commonly used to create epitaxial GaP layers on Si [9]. However, this technique makes it difficult to control the doping profile and inevitably reduces the service life of the shutters in the molecular beam epitaxy (MBE) system [10]. It is for these reasons that extensive research into alternative approaches to growing GaP buffer layers on Si is still underway [11].

The goal of this study is to develop a two-stage method of epitaxial growth of single-domain GaP buffer layers on a silicon substrate, which would ensure the separation of nucleation and growth stages of the gallium phosphide layer and include both low- and high-temperature stages, additionally allowing to forego the migration-enhanced epitaxy method.

For this purpose, it was necessary to establish the main factors determining the orientation of GaP lattice during nucleation on the vicinal surface of Si (001).

Two-stage method of epitaxial growth of GaP on Si (001)

Epitaxial heterostructures were grown on three-inch (75 mm) silicon wafers by molecular beam epitaxy (MBE). The heterostructures were grown on *p*-type vicinal Si (001) substrates with the electrical resistivity ranging from 0.3 to 3.0 Ohms·cm and the misorientation angle equal to $4.0 \pm 0.5^{\circ}$ in the <110> direction.

The Veeco GEN-III PA-MBE (plasma-assisted molecular beam epitaxy) system (Veeco, USA), equipped with cracker-type phosphorus and arsenic sources with needle valves was used in the study. The surface temperature was controlled by a thermocouple and a pyrometer calibrated by monitoring the surface phase transition Si (111) $7 \times 7 \rightarrow 1 \times 1$. The sources were calibrated by recording the molecular beam pressure ($P(P_2)$, P(Ga), P(Al)) with a Bayard–Alpert ionization gauge. The stoichiometric flux ratio P_2/Ga (pressure ratio $P(P_2)/P(\text{Ga}) \approx 6$) was found by monitoring the accumulation of gallium droplets on the surface of the structure.

Before the silicon wafers were loaded into the MBE chamber, they were cleaned by the Shiraki method; a thin layer of SiO_x oxide was formed on the surface of the wafer at the final stage of cleaning [12]. The surfaces of the Si (001) substrate and the forming epitaxial layer of gallium phosphide were studied in *situ* by reflection high-energy electron diffraction (RHEED). The surface silicon oxide was removed by thermal annealing at $T = 820 \pm 10$ °C (minimum required) for 30 minutes. RHEED patterns exhibited one type of (2×1) reconstructed Si (001) surface.

Notably, higher annealing temperatures contributed to a (2 × 2) surface reconstruction. The detected phenomenon is likely associated with contamination of the silicon surface due to an

increase in the background pressure of group V elements [13, 14]. In order to reduce this pressure, relatively low temperatures were used for the sources, so that low growth rates were consequently obtained. The growth rate of the gallium phosphide layer was 90 nm/hour.

A uniform distribution of molecular fluxes during the growth process was achieved by azimuthal rotation of the substrate at 5 rpm. The phosphorus flux value P_2 used for growing GaP and AlGaP layers exceeded the stoichiometric value by 1.33 times (pressure ratio $P(V)/P(III) = 8$).

An increase in the surface density and lateral size of defects representing voids with the depth exceeding the thickness of the formed layer was observed for a gallium phosphide layer grown with a stoichiometric ratio of P_2/Ga fluxes. An excessive number of Ga atoms apparently accumulated at the initial stages of growth at low values of the $_{P2}$ flux, serving to etch the silicon surface. The surface roughness increased significantly with a further increase in the ratio of P_2/Ga fluxes, up to values exceeding the stoichiometric value by half or more, as a result of suppressed diffusion of group III atoms.

The technique we propose for epitaxial growth of gallium phosphide buffer layers on silicon substrates includes two stages: low- and high-temperature. The seed AlGaP layer and the first GaP layer are grown at the first stage, and the second GaP layer is grown at the second stage. This technology provides excellent conditions for a heterostructure to form both at the nucleation stage and at the growth stage of the epitaxial layer.

The main stages of the growth process and the structural model of the buffer layer are shown in Fig. 1.

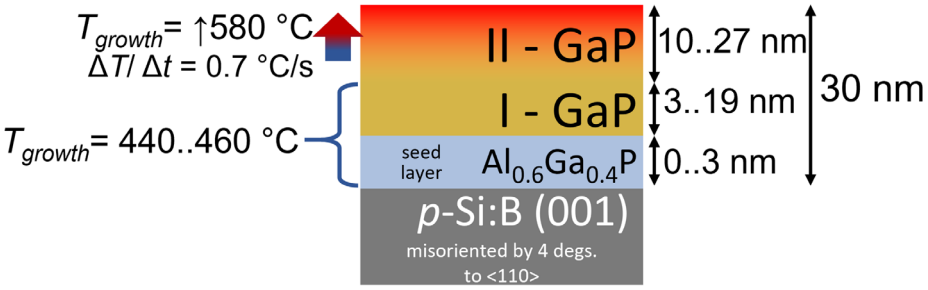

Fig. 1. Main steps of two-stage technique for growing GaP structure on *p*-Si(001) substrate (data on the left) and general structural model (the thickness values of low- and high-temperature GaP layers are given on the right)

We previously established that this technique allows to produce layers with smooth surface morphology [15], but studies of nucleation and growth of the crystalline structure of the layers went beyond the scope of the analysis. A separate low-temperature stage was considered in this study, with a seed AlGaP layer introduced, because a homogeneous epitaxial layer had to be grown over the entire surface of the substrate [16] at the initial stage. Therefore, immediately after the surface oxide was removed, the silicon substrate was cooled to a sufficiently low growth temperature (440–460 °C), ensuring a small diffusion length of adatoms (atoms on the crystal surface) and a high nucleation density of GaP islands while preserving their structural quality. The substrate was exposed to a P_2 flux for 10 seconds immediately before the start of the growth process.

The growth process was not interrupted during transition to the high-temperature stage, and the substrate temperature was raised to 580 °C at a rate of 0.7 °C/s. Conversely, experiments indicate that if the growth process was stopped, the layer was no longer homogeneous, with voids appearing on its surface, which persisted and were not overgrown upon further growth of the gallium phosphide layer.

The temperature providing high quality of GaP (001) layers during homoepitaxy as chosen based on literature data [17] for the high-temperature stage. The high mobility of adatoms contributed to a decrease in the density of structural defects, smoothing of the growth surface and consequent surface reconstruction observed from RHEED patterns. The total thickness of the GaP layer did not exceed 30 nm to avoid the formation of structural defects caused by inelastic relaxation of misfit stresses. The optimal conditions for growing gallium phosphide on silicon

substrate were determined by synthesizing samples with different temperatures chosen for forming the low-temperature I–GaP layer (440 and 460 ℃), different thicknesses of this layer (from 3 to 19 nm) and the seed layer that was a solid $Al_{0.6}Ga_{0.4}P$ solution (from 0 to 3 nm).

The quality of the buffer layer was preliminarily evaluated by atomic force microscopy (AFM) in tapping mode by comparing the roughness of the surface formed during coalescence of anti-phase domains. The structural characteristics of buffer layers on silicon were studied by transmission electron microscopy (TEM) and high-resolution X-ray diffractometry (HRXRD) in symmetric and asymmetric scattering geometries; this made it possible to establish the lattice matching in the plane of the layer and estimate the elastic stresses. The Panalytical X'PERT PRO MRD Extended X-ray diffractometer (Panalytical, the Netherlands) was used for HRXRD studies.

Growth of gallium phosphide on a silicon substrate

As known from the literature, two mutually orthogonal orientations of the crystal lattice with oppositely ordered sublattices of group III and V atoms can be observed under heteroepitaxy of $A^{III}B^{V}$ compounds on the silicon surface, due to lower symmetry of the crystal lattice of the epitaxial layer compared with the substrate [9, 10]. For this reason, a crucial condition for growing single-domain GaP on Si(001) is the stabilization of the homogeneous P-terminated (2×1) Si(001) surface. Furthermore, it is required to find the nucleation conditions under which one type of chemical bond is established over the entire area of the heterointerface.

RHEED measurements allowed to not only study the state of the silicon surface but also to determine the preferred orientation of the nucleating GaP layer. Varying the incidence angle of the RHEED beam relative to the (001) crystallographic plane and, consequently, varying the position of the Bragg reflections and Kikuchi bands observed under azimuthal rotation of the vicinal substrate allowed to establish the orientation of the surface reconstructions of Si and GaP relative to the azimuthal misorientation of the substrate (Fig. 2).

Evidently, after thermal annealing of vicinal silicon substrate, superstructural reflections with double periodicity are only observed for the electron beam incident along the direction of azimuthal misorientation. This observation indicates that the dimer rows forming one type of (2×1) reconstructed Si surface are predominantly oriented perpendicular to the edges of the atomic

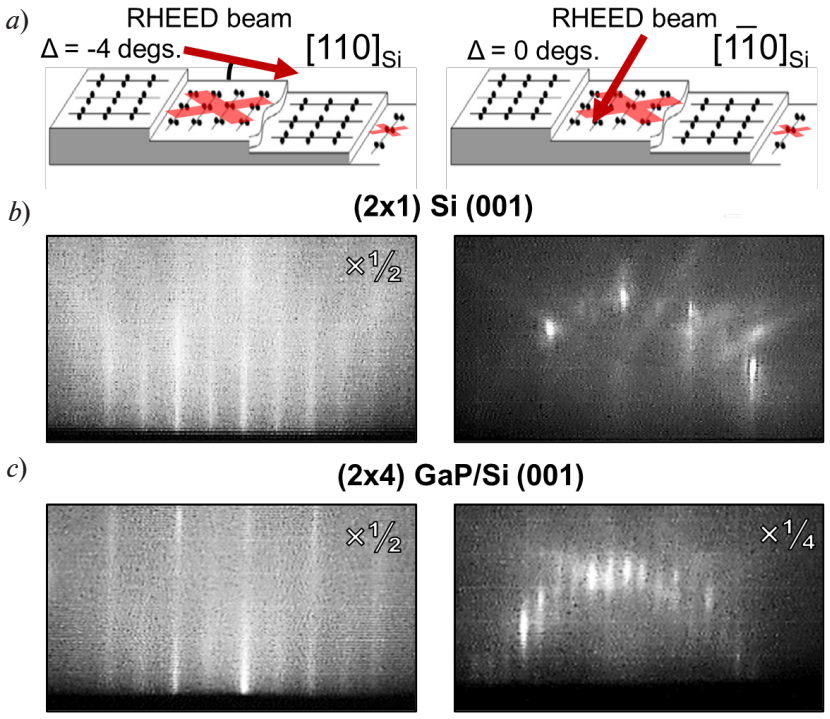

Fig. 2. Schematic for RHEED measurements in mutually orthogonal azimuthal orientations of Si substrate (a); resulting RHEED patterns for Si (001) (2×1) (b) and GaP (001) (2×4) (c) surface reconstructions

steps and, consequently, most of the surface area of the substrate is occupied by atomic terraces separated by S_B monoatomic steps or D_B double steps. It was previously reported [10, 18–20] that one type of (2 × 1) reconstructed Si surface can be observed at high temperature (1000 °C) annealing of the vicinal Si (001) surface, which leads to the formation of a system of double atomic steps. We assume that an orthogonally oriented Si (1 × 2) reconstruction was not observed due to narrow atomic terraces on the vicinal Si surface (less than 10 nm) and low annealing temperature (820 °C), insufficient for activation of surface diffusion in silicon and rearrangement of the system of atomic steps [21].

Analyzing the RHEED patterns for the high-temperature growth stage, we concluded that, two different orientations are obtained for the P-stabilized (2 × 4) reconstructed GaP (001) surface, depending on the temperature of the substrate selected for the low-temperature stage of GaP nucleation (440 or 460 °C).

The corresponding AFM images of the surface morphology and the RHEED patterns of GaP structures on Si with a 3 nm thick low-temperature I-GaP layer are shown in Fig. 3, *a* and *b*. In the case of GaP nucleation on a P-terminated silicon surface at 440 °C, the $[1\bar{1}0]$ direction [1/20] of GaP lattice turns out to be oriented perpendicular to dimer rows (parallel to the edges of the atomic steps in Si) forming the Si (2 × 1) superstructure [17]. Evidently, the Si (2 × 1) surface reconstruction corresponds to the GaP (2 × 4) surface reconstruction (Fig. 2,*c*).

The azimuthal orientations of RHEED patterns at which superstructural reflections are observed for GaP (2 × 4) are rotated by 90° in the case of GaP nucleation on Si surface at 460 °C (see Fig. 3,*b*); in this case, the Si (2 × 1) surface reconstruction corresponds to the GaP (4 × 2) reconstruction. Notably, rotation of the azimuthal orientation of GaP surface reconstruction cannot be caused by the transition from a phosphorus-stabilized to gallium-stabilized surface, as observed, for example, for the GaAs(001) surface [22], since both types of GaP (001) surface termination have identical periodicity and orientation of the (2 × 4) surface superstructure [23]. It is most likely that not only termination but also substitution of silicon atoms with phosphorus atoms becomes possible at higher temperatures of Si (001) surface. As a result, the $[1\bar{1}0]$ direction of the GaP lattice is oriented parallel to Si dimer rows (perpendicular to the edges of Si atomic steps), and the Ga and P sublattices turn out to be inverted relative to the first case.

It can be seen from the AFM images that the surface morphology of the samples also varies depending on the temperature selected at the initial growth stage. If the growth temperature of

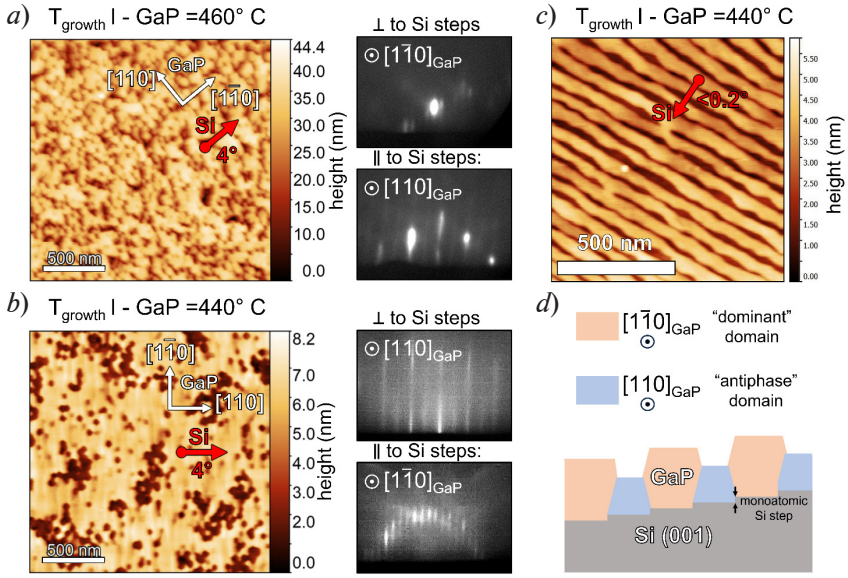

Fig. 3. Morphology and crystallographic orientation of GaP layers on vicinal Si (001) substrates (a–c); scheme for overgrowth of antiphase regions by dominant domain (d). AFM surface images and corresponding RHEED patterns (collected at the final stage of layer growth) for growth temperatures of 460 °C (a) and 440 °C (b); AFM image for misorientation angle of 0.2 ° (c). The red arrows indicate the directions of azimuthal misorientation of the substrate

the I-GaP layer is 440 °C, a relatively smooth surface is observed with further high-temperature growth of GaP, embedded with separate regions deepened by 5–6 nm. The surface of the GaP layer formed on the low-temperature I-GaP layer grown at 460 °C consists of separate islands varying in height by about 30 nm.

The diffusion length of adatoms and the energy of the Ehrlich–Schwoebel barrier on the surface of $A^{III}B^{V}$ (001) compounds depend on the mutual orientation of step edges and the number of broken chemical bonds on the surface [24–26]. Therefore, the growth rate of GaP islands becomes higher if the [1$\bar{1}$0] direction of the GaP lattice turns out to be oriented parallel to the edge of the atomic steps on the growth surface.

This effect was demonstrated for a structure grown on Si (001) surface with a small misorientation angle (about 0.2°) towards the <110> azimuth. The corresponding AFM image of the structure surface is shown in Fig. 3,*c*). Due to the anisotropy of growth rates, a prominent GaP domain is formed on the surface of the structure, and its islands evolve along the [1$\bar{1}$0] direction oriented parallel to the edges of the atomic steps (see the schematic in Fig. 3,*d*).

As a result, atomically smooth regions are observed in the AFM images, whose surface rises above the antiphase regions of the layer with an orthogonally oriented GaP lattice (dark regions in AFM images, see Fig. 3.). Appearance of the corresponding antiphase regions in GaP can be associated both with the presence of atomic steps on the silicon surface with orthogonally oriented broken silicon bonds, and with the inhomogeneous chemical bonding at the GaP/Si interface. The first scenario is illustrated in Fig. 3,*c*: this is the case of the silicon surface with a small misorientation angle.

Thus, the growth experiments conducted indicate that the formation of antiphase domains is effectively suppressed only for a single mutual orientation of GaP lattice relative to the edges of silicon steps, observed under low-temperature nucleation of GaP on Si (001) surface at 440 °C. It was found that the proportion of the surface occupied by antiphase regions decreases with increasing thickness of the low-temperature I-GaP layer. The corresponding AFM image of surface morphology of the GaP layer grown using the 10 nm thick I-GaP layer is shown in Fig. 4,*a*. The short diffusion length and high nucleation density at the low-temperature growth stage contribute to effective overgrowth of antiphase domains at the initial growth stage.

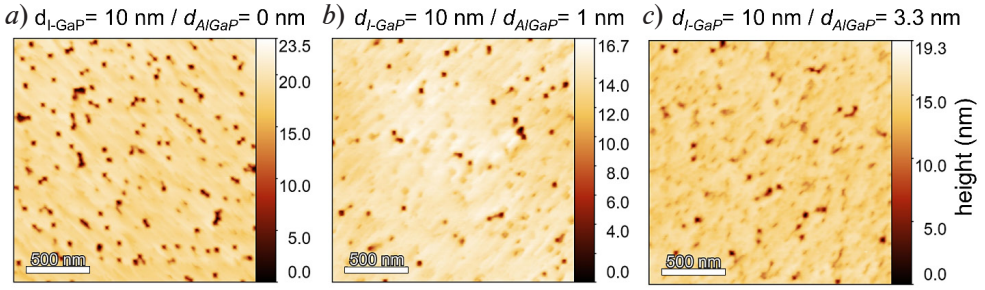

Fig. 4. AFM images for surface morphology of I-GaP layers (10 nm thick) on Si (001), illustrating the influence of $Al_{0.6}Ga_{0.4}P$ seed layer thicknesses, nm: 0.0 (a), 1.0 (b), 3.3 (c)

However, the height difference between the surfaces of the dominant and antiphase domains also increases due to the small lateral size (less than 10 nm) and the low growth rate of the antiphase regions. Consequently, void-like defects which not overgrown at the high-temperature stage may form on the surface of the GaP layer.

It was found in [27] that the presence of chemically active aluminum adatoms at the initial growth stage ensures uniform nucleation of the AlGaP solid solution layer and helps prevent an interfacial chemical reaction between Ga and Si.

We conducted a series of growth experiments, discovering that pre-growing a seed layer of $Al_{0.6}Ga_{0.4}P$ or AlP solid solution no more than 1 nm thick at the low-temperature stage improves the smoothness of the surface of GaP layers. A comparative analysis for the dependence of buffer layer morphology on the thickness of the AlGaP seed layer is shown in Fig. 4. It is most likely that Al adatoms have a shorter diffusion length due to the higher chemical bond energy, which contributes to more homogeneous nucleation of the AlGaP layer on Si compared with GaP.

Conversely, excessively high thicknesses of the AlGaP seed layer (3 nm or more) produce a rougher growth surface.

The described observations show that overgrowth of antiphase regions in the GaP layer should be carried out at the low-temperature stage. The thickness of the low-temperature layer should be comparable with the lateral size of the antiphase regions. The subsequent high-temperature growth stage determines the structural perfection and atomic smoothness of the GaP surface.

Therefore, it is possible to almost completely suppress the formation of antiphase regions as well as increase the smoothness of the GaP layer surface, if an AlGaP seed layer about 1 nm thick is used, and the thickness at the low-temperature stage is doubled, up to 20 nm. The final arrangement of the structure layers as well as the AFM image of the surface are shown in Fig. 5, *a* and *b*, respectively.

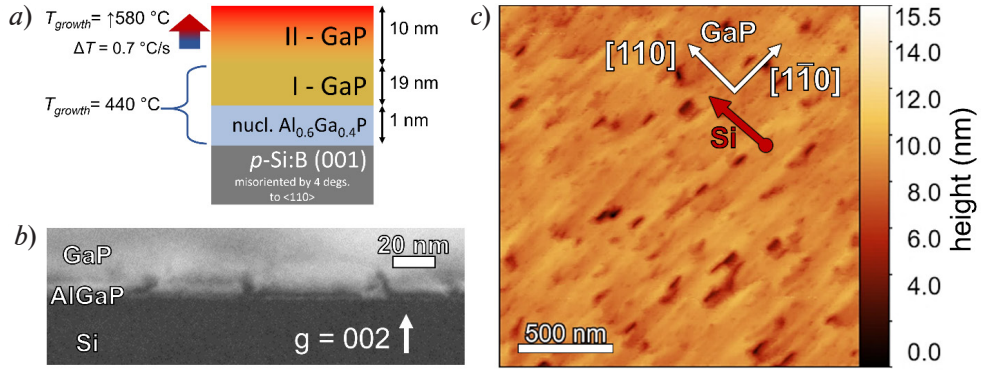

Fig. 5. Experimental search for optimal conditions for single-domain formation of GaP buffer layer on Si(001) with atomically smooth surface: model of resulting structure with growth factors (*a*); AFM image of surface morphology (*b*) (main crystallographic directions of GaP and direction of azimuthal orientation of Si are marked by white and red arrows, respectively); dark-field TEM image with diffraction contrast (*c*) (**g** = 002 is the applied diffraction vector)

Structural properties of thin GaP layers on Si (001)

The microscopic nature of structural defects in the GaP buffer layers on Si grown by the optimized two-stage technique was investigated by the TEM method. A characteristic dark-field TEM image with diffraction contrast for orientations of the diffraction vector **g** = 002 is shown in Fig. 5,*c*. A thin layer formed by a network of misfit dislocations is visible in the region of the heterointerface with the substrate [28]. Notably, the contrast associated with antiphase domains could not be reliably detected in dark-field TEM images [29].

Accordingly, the TEM data confirm that the formation of antiphase domains stops at the initial stages of buffer layer growth.

Integral information on the structural characteristics of GaP buffer layers was obtained by HRXRD in symmetric and asymmetric scattering geometries; this allowed to establish the lattice matching in the plane of the layer and to estimate the elastic stresses. The XRD ω–2θ scans, obtained for the GaP layer grown by the optimized two-stage technique (Fig. 6,*a*),exhibit well-resolved interference fringes pointing to the formation of an atomically smooth GaP surface and a sharp GaP/Si heterointerface.

The period of these fringes corresponds to a GaP layer thickness of 33 nm, which is in good agreement with the computational value (30 nm). The reciprocal space map obtained for the diffraction intensity near the asymmetric Si (224) Bragg peak (see Fig. 6,*b*) indicates that the GaP lattice in the (001) plane exactly repeats the substrate lattice, i.e., Bragg reflections of the film and the substrate correspond to identical values of lateral momentum transfer Q_{xy}.

To summarize, uniaxial anisotropic deformation along the [001] direction is observed in the GaP layer. In turn, the ω-rocking curves near the (004) Bragg peak (see the inset in Fig. 6,*a*) exhibit both an intense narrow coherent component with a FWHM of about 11–12″, which is

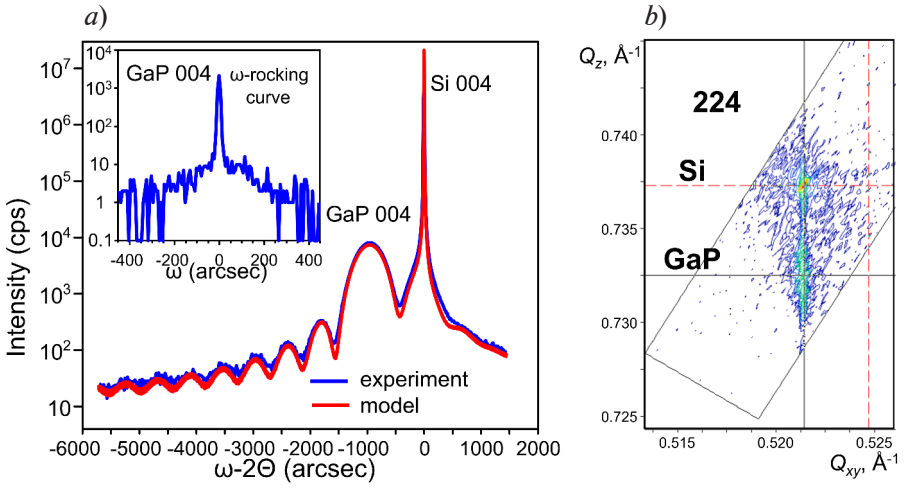

Fig. 6. Data from high-resolution X-ray diffractometry of GaP buffer layers on Si substrate: experimental (blue) and computational (red) X-ray diffraction ω–2θ scans near specular Si (004) Bragg peak (*a*); reciprocal space map of diffraction intensity distribution near asymmetric Si (224) Bragg peak (*b*) Inset: ω-rocking curve near GaP (004) reflection

only twice the FWHM of the substrate peak, and a weak wide diffuse component (more than 200″), which indicates a nonuniform distribution of elastic stresses. In general, the rocking curves indicate a high degree of lateral spatial coherence of the epitaxial layer (about 2.45 μm), suggesting a high level of structural perfection.

Conclusion

In this paper, we established the influence of growth factors and the state of the silicon surface on the formation of GaP buffer layers on Si(001) substrates.

It was proved that single-domain GaP layers of high structural perfection can be grown epitaxially on the vicinal Si (001) surface with a misorientation angle and azimuth of 4.0 ± 0.5° and <110>, respectively, without resorting to migration-enhanced epitaxy methods or synthesis of homoepitaxial buffer layers of silicon.

We also clarified the effect of seed layers in solid AlGaP structures of nanometer thickness on the process of annihilation of antiphase domains and the density of point defects on the GaP surface.

A two-stage procedure was developed for epitaxial growth of GaP buffer layers on Si (001), providing separation of the stages of nucleation and growth of the GaP layer and consisting of low-temperature (440 °C) and high-temperature (580 °C) stages. Unlike the migration-enhanced epitaxy method, the proposed approach provides the ability to control the doping profile of GaP buffer layers on Si. The thin GaP buffer layers grown on silicon have a pseudomorphic single-domain structure and an atomically smooth surface (the RMS roughness was less than 2 nm).

We believe that our study makes as an important step towards further advances in the technology for growing $A^{III}B^{V}$ heterostructures on silicon, valuable for manufacturing applications.

REFERENCES

1. **Roelkens G., Liu L., Liang D., et al.,** III-V/silicon photonics for on-chip and intra-chip optical interconnects: III-V/silicon photonics for optical interconnects, Laser Photonics Revs. 4 (6) (2010) 751–779.

2. **Lee K. E., Fitzgerald E. A.,** Metamorphic transistors: Building blocks for hetero-integrated circuits, MRS Bull. 41 (3) (2016) 210–217.

3. **Richardson C. J. K., Lee M. L.,** Metamorphic epitaxial materials // MRS Bull. 41 (3) (2016) 193–198.

4. **Takano Y., Hisaka M., Fugii N., et al.,** Reduction of threading dislocations by InGaAs interlayer in GaAs layers grown on Si substrates, Appl. Phys. Lett. Letters. 73 (20) (1998) 2917–2919.

5. **Wang Y. P., Letoublon A., Thangh T. N., et al.,** Quantitative evaluation of microtwins and antiphase defects in GaP/Si nanolayers for a III–V photonics platform on silicon using a laboratory X-ray diffraction setup, J. Appl. Crystallogr. 48 (3) (2015) 702–710.

6. **Vitali L., Ramsey M. G., Netzer F. C.,** Substitutional geometry and strain effects in overlayers of phosphorus on Si(111), Phys. Rev. B. 57 (24) (1998) 15376–15384.

7. **Ji J.-Y., Shen T.-C.,** A scanning tunneling microscopy study of PH_3 adsorption on Si(111)-7×7 surfaces, P-segregation and thermal desorption, Surf. Sci. 601 (7) (2007) 1768–1774.

8. **Khoury M., Tottereau O., Feuillet G., et al.,** Evolution and prevention of meltback etching: Case study of semipolar GaN growth on patterned silicon substrates, J. Appl. Phys. 122 (10) (2017) 105108.

9. **Takagi Y., Yonezu H., Samonji K., et al.,** Generation and suppression process of crystalline defects in GaP layers grown on misoriented Si(100) substrates, J. Cryst. Growth. 187 (1) (1998) 42–50.

10. **Volz K., Beyer A., Witte W., et al.,** GaP-nucleation on exact Si (001) substrates for III/V device integration, J. Cryst. Growth. 315 (1) (2011) 37–47.

11. **Yamane K., Kobayashi T., Furukawa Y., et al.,** Growth of pit-free GaP on Si by suppression of a surface reaction at an initial growth stage, J. Cryst. Growth. 311 (3) (2009) 794–797.

12. **Okumura H., Akane T., Tsubo Y., Matsumoto S.,** Comparison of conventional surface cleaning methods for Si molecular beam epitaxy, J. Electrochem. Soc. 144 (11) (1997) 3765–3768.

13. **Curson N. J., Schofield S. R., Simmons M. Y., et al.,** STM characterization of the Si-P heterodimer, Phys. Rev. B. 69 (19) (2004) 195303.

14. **Zhang C., Kim Y., Faleev N. N., Honsberg C. B.,** Improvement of GaP crystal quality and silicon bulk lifetime in GaP/Si heteroepitaxy, J. Cryst. Growth. 475 (1 Oct) (2017) 83–87.

15. **Sapunov G. A., Koval A. Yu., Fedorov V. V., Bolshakov A. D.,** Epitaxial synthesis of single-domain gallium phosphide on silicon, J. Phys. Conf. Ser. 1697 (2020) 012127.

16. **Narayanan V., Mahajan S., Bachmann K. J., et al.,** Stacking faults and twins in gallium phosphide layers grown on silicon, Philos. Mag. A. 82 (4) (2002) 685–698.

17. **Ratcliff C., Grassman T. J., Carlin J. A., Ringel S. A.,** High temperature step-flow growth of gallium phosphide by molecular beam epitaxy and metalorganic chemical vapor deposition, Appl. Phys. Lett. 99 (14) (2011) 141905.

18. **Martin M., Caliste D., Cipro R., et al.,** Toward the III-V/Si co-integration by controlling the biatomic steps on hydrogenated Si(001), Appl. Phys. Lett. 109 (25) (2016) 253103.

19. **Wierenga P. E., Kubby J. A., Griffith J. E.,** Tunneling images of biatomic steps on Si(001), Phys. Rev. Lett. 59 (19) (1987) 2169–2172.

20. **Soga T., Jimbo T., Umeno M.,** Dislocation generation mechanisms for GaP on Si grown by metalorganic chemical vapor deposition, Appl. Phys. Lett. 63 (18) (1993) 2543–2545.

21. **Persichetti L., Capasso A., Sgariata A., et al.,** Towards a controlled growth of self-assembled nanostructures: Shaping, ordering, and localization in Ge/Si heteroepitaxy (Ch.), In book: S. Bellucci (Ed.), "Self-Assembly of Nanostructures", Book Ser. "Lecture Notes in Nanoscale Science and Technology", Vol. 12. Springer, New York (2012) 201–263.

22. **Schmidt W. G.,** III-V compound semiconductor (001) surfaces, Appl. Phys. A: Mater. Sci. Process. 75 (1) (2002) 89–99.

23. **Baillargeon J. N., Cheng K. Y., Hsieh K. C.,** Surface structure of (100) GaP grown by gas source molecular beam epitaxy, Appl. Phys. Lett. 56 (22) (1990) 2201–2203.

24. **Beyer A., Ohlmann J., Liebich S., et al.,** GaP heteroepitaxy on Si(001): Correlation of Si-surface structure, GaP growth conditions, and Si-III/V interface structure, J. Appl. Phys. 111 (8) (2012) 083534.

25. **Naresh-Kumar G., Vilalta-Clemente A., Jussila H., et al.,** Quantitative imaging of anti-phase domains by polarity sensitive orientation mapping using electron backscatter diffraction, Sci. Rep. (07 Sept) (2017) 10916.

26. **Lin A. C., Fejer M. M., Harris J. S.,** Antiphase domain annihilation during growth of GaP on Si by molecular beam epitaxy, J. Cryst. Growth. 363 (15 Jan) (2013) 258–263.

27. **Lin A. C., Bassiri R., Omar S., et al.,** Epitaxial growth of GaP/AlGaP mirrors on Si for low thermal noise optical coatings, Opt. Mater. Express. 5 (8) (2015) 1890–1897.

28. **Pantzas K., Beaudoin G., Bailly M., et al.,** Relaxation mechanism of GaP grown on 001 Si substrates: influence of defects on the growth of AlGaP layers on GaP/Si templates, Philos. Mag. 101 (20) (2021) 2189–2199.

29. **Narayanan V., Mahajan S., Bachmann K. J., et al.,** Antiphase boundaries in GaP layers grown on (001) Si by chemical beam epitaxy, Acta Mater. 50 (6) (2002) 1275–1287.

СПИСОК ЛИТЕРАТУРЫ

1. **Roelkens G., Liu L., Liang D., Jones R., Fang A., Koch B., Bowers J.** III-V/silicon photonics for on-chip and intra-chip optical interconnects: III-V/silicon photonics for optical interconnects // Laser & Photonics Reviews. 2010. Vol. 4. No. 6. Pp. 751–779.

2. **Lee K. E., Fitzgerald E. A.** Metamorphic transistors: Building blocks for hetero-integrated circuits // MRS (Materials Research Society) Bulletin. 2016. Vol. 41. No. 3. Pp. 210–217.

3. **Richardson C. J. K., Lee M. L.** Metamorphic epitaxial materials // MRS Bulletin. 2016. Vol. 41. No. 3. Pp. 193–198.

4. **Takano Y., Hisaka M., Fugii N., Suzuki K., Kuwahara K., Fuke S.** Reduction of threading dislocations by InGaAs interlayer in GaAs layers grown on Si substrates // Applied Physics Letters. 1998. Vol. 73. No. 20. Pp. 2917–2919.

5. **Wang Y. P., Letoublon A., Thangh T. N., Bahri M., Largeau L., Patriarche G., Cornet C., Bertru N., Le Corre A., Durand O.** Quantitative evaluation of microtwins and antiphase defects in GaP/Si nanolayers for a III–V photonics platform on silicon using a laboratory X-ray diffraction setup // Journal of Applied Crystallography. 2015. Vol. 48. No. 3. Pp. 702–710.

6. **Vitali L., Ramsey M. G., Netzer F. C.** Substitutional geometry and strain effects in overlayers of phosphorus on Si(111) // Physical Review B. 1998. Vol. 57. No. 24. Pp. 15376–15384.

7. **Ji J.-Y., Shen T.-C.** A scanning tunneling microscopy study of PH_3 adsorption on Si(111)-7×7 surfaces, P-segregation and thermal desorption // Surface Science. 2007. Vol. 601. No. 7. Pp. 1768–1774.

8. **Khoury M., Tottereau O., Feuillet G., Vennégués P., Zúciga-Pérez J.** Evolution and prevention of meltback etching: Case study of semipolar GaN growth on patterned silicon substrates // Journal of Applied Physics. 2017. Vol. 122. No. 10. P. 105108.

9. **Takagi Y., Yonezu H., Samonji K., Tsuji T., Ohshima N.** Generation and suppression process of crystalline defects in GaP layers grown on misoriented Si(100) substrates // Journal of Crystal Growth. 1998. Vol. 187. No. 1. Pp. 42–50.

10. **Volz K., Beyer A., Witte W., Ohlmann J., Németh I., Kunert B., Stolz W.** GaP-nucleation on exact Si (001) substrates for III/V device integration // Journal of Crystal Growth. 2011. Vol. 315. No. 1. Pp. 37–47.

11. **Yamane K., Kobayashi T., Furukawa Y., Okada H., Yonezu H., Wakahara A.** Growth of pit-free GaP on Si by suppression of a surface reaction at an initial growth stage // Journal of Crystal Growth. 2009. Vol. 311. No. 3. Pp. 794–797.

12. **Okumura H., Akane T., Tsubo Y., Matsumoto S.** Comparison of conventional surface cleaning methods for Si molecular beam epitaxy // Journal of The Electrochemical Society. 1997. Vol. 144. No. 11. Pp. 3765–3768.

13. **Curson N. J., Schofield S. R., Simmons M. Y., Oberbeck L., O'Brien J. L., Clark R. G.** STM characterization of the Si-P heterodimer // Physical Review B. 2004. Vol. 69. No. 19. P. 195303.

14. **Zhang C., Kim Y., Faleev N. N., Honsberg C. B.** Improvement of GaP crystal quality and silicon bulk lifetime in GaP/Si heteroepitaxy // Journal of Crystal Growth. 2017. Vol. 475. 1 October. Pp. 83–87.

15. **Sapunov G. A., Koval A. Yu., Fedorov V. V., Bolshakov A. D.** Epitaxial synthesis of single-domain gallium phosphide on silicon // Journal of Physics: Conference Series. 2020. Vol. 1697. P. 012127.

16. **Narayanan V., Mahajan S., Bachmann K. J., Woods V., Dietz N.** Stacking faults and twins in gallium phosphide layers grown on silicon // Philosophical Magazine A. 2002. Vol. 82. No. 4. Pp. 685–698.

17. **Ratcliff C., Grassman T. J., Carlin J. A., Ringel S. A.** High temperature step-flow growth of gallium phosphide by molecular beam epitaxy and metalorganic chemical vapor deposition // Applied Physics Letters. 2011. Vol. 99. No. 14. P. 141905.

18. **Martin M., Caliste D., Cipro R., et al.** Toward the III-V/Si co-integration by controlling the biatomic steps on hydrogenated Si(001) // Applied Physics Letters. 2016. Vol. 109. No. 25. P. 253103.

19. **Wierenga P. E., Kubby J. A., Griffith J. E.** Tunneling images of biatomic steps on Si(001) // Physical Review Letters. 1987. Vol. 59. No. 19. Pp. 2169–2172.

20. **Soga T., Jimbo T., Umeno M.** Dislocation generation mechanisms for GaP on Si grown by metalorganic chemical vapor deposition // Applied Physics Letters. 1993. Vol. 63. No. 18. Pp. 2543–2545.

21. **Persichetti L., Capasso A., Sgariata A., Fanfoni M., Motta N., Balzarotti A.** Towards a controlled growth of self-assembled nanostructures: Shaping, ordering, and localization in Ge/Si heteroepitaxy (Chapter) // S. Bellucci (Editor). "Self-Assembly of Nanostructures". Book Series: Lecture Notes in Nanoscale Science and Technology. Vol. 12. New York: Springer, 2012. Pp. 201–263.

22. **Schmidt W. G.** III-V compound semiconductor (001) surfaces // Applied Physics A: Materials Science and Processing. 2002. Vol. 75. No. 1. Pp. 89–99.

23. **Baillargeon J. N., Cheng K. Y., Hsieh K. C.** Surface structure of (100) GaP grown by gas source molecular beam epitaxy // Applied Physics Letters. 1990. Vol. 56. No. 22. Pp. 2201–2203.

24. **Beyer A., Ohlmann J., Liebich S., Heim H., Witte G., Stolz W., Volz K.** GaP heteroepitaxy on Si(001): Correlation of Si-surface structure, GaP growth conditions, and Si-III/V interface structure // Journal of Applied Physics. 2012. Vol. 111. No. 8. P. 083534.

25. **Naresh-Kumar G., Vilalta-Clemente A., Jussila H., Winkelmann A., Nolze G., Vespucci S., Nagarajan S., Wilkinson A. J., Trager-Cowan C.** Quantitative imaging of anti-phase domains by polarity sensitive orientation mapping using electron backscatter diffraction // Scientific Reports. 2017. 07 September. P. 10916.

26. **Lin A. C., Fejer M. M., Harris J. S.** Antiphase domain annihilation during growth of GaP on Si by molecular beam epitaxy // Journal of Crystal Growth. 2013. Vol. 363. 15 January. Pp. 258–263.

27. **Lin A. C., Bassiri R., Omar S., Markosyan A. S., Lanz B., Route R., Byer R.L., Harris J. S., Fejer M. M.** Epitaxial growth of GaP/AlGaP mirrors on Si for low thermal noise optical coatings // Optical Materials Express. 2015. Vol. 5. No. 8. Pp. 1890–1897.

28. **Pantzas K., Beaudoin G., Bailly M., Martin A., Grisard A., Dolfi D., Mauguin O., Largeau L., Sagnes I., Patriarche G.** Relaxation mechanism of GaP grown on 001 Si substrates: influence of defects on the growth of AlGaP layers on GaP/Si templates // Philosophical Magazine. 2021. Vol. 101. No. 20. Pp. 2189–2199.

29. **Narayanan V., Mahajan S., Bachmann K. J., Woods V., Dietz N.** Antiphase boundaries in GaP layers grown on (001) Si by chemical beam epitaxy // Acta Materialia. 2002. Vol. 50. No. 6. Pp. 1275–1287.

THE AUTHORS

FEDOROV Vladimir V.
Peter the Great St. Petersburg Polytechnic University,
Alferov University, RAS
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
fedorov_vv@spbstu.ru
ORCID: 0000-0001-5547-9387

FEDINA Sergey V.
Peter the Great St. Petersburg Polytechnic University,
Alferov University, RAS
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
fedina.serg@yandex.ru
ORCID: 0000-0001-7521-3754

KAVEEV Andrey K.
Ioffe Institute, RAS,
Alferov University, RAS
26, Polytekhnicheskaya St., St. Petersburg, 194021, Russia
kaveev@mail.ioffe.ru
ORCID: 0000-0002-3640-677X

KIRILENKO Demid A.
Ioffe Institute, RAS,
ITMO University
26, Polytekhnicheskaya St., St. Petersburg, 194021, Russia
zumsisai@gmail.com
ORCID: 0000-0002-1571-209X

FALEEV Nikolai N.
Ioffe Institute, RAS
26, Polytekhnicheskaya St., St. Petersburg, 194021, Russia
nfaleev@mail.ioffe.ru

MUKHIN Ivan S.
Peter the Great St. Petersburg Polytechnic University,
Alferov University, RAS
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
imukhin@yandex.ru
ORCID: 0000-0001-9792-045X

СВЕДЕНИЯ ОБ АВТОРАХ

ФЕДОРОВ Владимир Викторович — *кандидат физико-математических наук, старший научный сотрудник Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого; старший научный сотрудник лаборатории возобновляемых источников энергии Академического университета им. Ж. И. Алфёрова РАН.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
fedorov_vv@spbstu.ru
ORCID: 0000-0001-5547-9387

ФЕДИНА Сергей Викторович — *инженер Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого, лаборант лаборатории возобновляемых источников энергии Академического университета им. Ж. И. Алфёрова РАН.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
fedina.serg@yandex.ru
ORCID: 0000-0001-7521-3754

КАВЕЕВ Андрей Камильевич — *доктор физико-математических наук, старший научный сотрудник лаборатории диагностики материалов и структур твердотельной электроники Физико-технического института им. А. Ф. Иоффе РАН, ведущий научный сотрудник лаборатории возобновляемых источников энергии Академического университета им. Ж. И. Алфёрова РАН.*
194021, г. Санкт-Петербург, Политехническая ул., 26
kaveev@mail.ioffe.ru
ORCID: 0000-0002-3640-677X

КИРИЛЕНКО Демид Александрович – *доктор физико- математических наук, PhD, старший научный сотрудник лаборатории диагностики материалов и структур твердотельной электроники Физико-технического института им. А. Ф. Иоффе РАН, доцент Института перспективных систем передачи данных Национального исследовательского университета информационных технологий, механики и оптики (ИТМО).*
194021, г. Санкт-Петербург, Политехническая ул., 26
zumsisai@gmail.com
ORCID: 0000-0002-1571-209X

ФАЛЕЕВ Николай Николаевич – *кандидат физико-математических наук, старший научный сотрудник лаборатории диагностики материалов и структур твердотельной электроники Физико-технического институт им. А. Ф. Иоффе РАН.*
194021, г. Санкт-Петербург, Политехническая ул., 26
nfaleev@mail.ioffe.ru

МУХИН Иван Сергеевич – *доктор физико-математических наук, директор Высшей инженерно-физической школы (ВИФШ), профессор ВИФШ Санкт-Петербургского политехнического университета Петра Великого; заведующий лабораторией возобновляемых источников энергии Академического университета им. Ж. И. Алфёрова РАН.*
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
imukhin@yandex.ru
ORCID: 0000-0001-9792-045X

Received 24.11.2023. Approved after reviewing 12.02.2024. Accepted 12.02.2024.
Статья поступила в редакцию 24.11.2023. Одобрена после рецензирования 12.02.2024. Принята 12.02.2024.

Original article
DOI: https://doi.org/10.18721/JPM.17211

THE TEMPERATURE EFFECT ON IMPULSE DIELECTRIC STRENGTH OF POLYMER FILMS

S. E. Semenov[1,2]✉

[1] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

[2] Ioffe Institute, St. Petersburg, Russia

✉ moritohayama96@gmail.com

Abstract. In the paper, the experimental data on the temperature effect on the breakdown strength of polypropylene, polycarbonate and poly(ethylene terephthalate) films 2.0–2.5 μm thick have been obtained over a temperature range from 293 to 363 K. When the film samples were heated, the breakdown electric field strength was found to decrease slightly but to reveal a significant scatter of values. It was shown that the experimental results on the pulsed electrical breakdown of the polymer films could be described basing the notion of the ionization mechanism of the polymer breakdown, not related to the development of an impact ionization of the polymer molecules.

Keywords: impulse dielectric strength, temperature, polypropylene, polycarbonate, poly(ethylene terephthalate)

Citation: Semenov S. E., The temperature effect on impulse dielectric strength of polymer films, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 17 (2) (2024) 141–149. DOI: https://doi.org/10.18721/JPM.17211

Научная статья
УДК 53.043
DOI: https://doi.org/10.18721/JPM.17211

ВЛИЯНИЕ ТЕМПЕРАТУРЫ НА ИМПУЛЬСНУЮ ЭЛЕКТРИЧЕСКУЮ ПРОЧНОСТЬ ПОЛИМЕРНЫХ ПЛЕНОК

С. Е. Семенов[1,2]✉

[1] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

[2] Физико-технический институт им. А. Ф. Иоффе РАН, Санкт-Петербург, Россия

✉ moritohayama96@gmail.com

Аннотация. В работе исследовано влияние температуры на пробивную напряженность пленок полиэтилена, поликарбоната и полиэтилентерефталата толщиной 2,5 – 2,0 мкм в температурном диапазоне 363 – 293 K. Установлено, что при нагреве образцов пленок пробивная напряженность электрического поля уменьшается незначительно, но обнаруживает существенный разброс значений. Показано, что экспериментальные данные по импульсному пробою полимерных пленок можно описать на основе представления об ионизационном механизме пробоя полимеров, не связанном с развитием ударной ионизации их молекул.

Ключевые слова: полимерная пленка, импульсный пробой, пробивная напряженность, полиэтилен, поликарбонат, полиэтилентерефталат, температура

Ссылка для цитирования: Семенов С. Е. Влияние температуры на импульсную электрическую прочность полимерных пленок // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2024. Т. 17. № 2. С. 141–149. DOI: https://doi.org/10.18721/JPM.17211

Introduction

Steady interest towards the dielectric strength of polymer dielectrics has persisted for many decades, due to the great scientific and practical implications of this characteristic for studies of the breakdown phenomenon. The vast majority of publications on this issue consider electrical breakdown in polymers exposed to DC and AC voltages. Considerably scarcer data are available on the picture of impulse breakdown of polymers, mainly consisting of results obtained decades ago [1–3]. Technical advances made in recent years allow detecting breakdown characteristics under nanosecond voltage pulses, recording breakdown times and voltages with high accuracy [4–9].

One of the most important factors influencing the evolution of electrical degradation and breakdown in polymer dielectrics is temperature. Its role under the influence of voltage pulses was studied in large samples of poly (methyl methacrylate) (PMMA) [2], polyethylene (PE) and polytetrafluoroethylene (PTFE) [3]. The effect of temperature on the breakdown strength F_{br} turned out to be different for different polymers. F_{br} decreased by more than 25% in PMMA samples under heating from 293 to 433 K, while the value of F_{br} for PE and PTFE remained virtually constant in the same temperature range.

Polymer films are a special group in the diverse range of dielectric materials. They are widely used in electrical engineering, for example, for manufacturing high-voltage capacitors, or for synthesis of films ranging in thickness from several to tens of μm that serve as convenient samples for studies of electrical aging and breakdown in polymer material. However, the effect of temperature on the impulse strength of polymer films is yet to be fully explored.

The goal of this study is to understand the nature of the temperature effect on the characteristics of impulse strength of various polymer films at room and elevated temperatures.

Experimental procedure

Industrial films made of polypropylene (PP) with the thickness of 2.0 μm, polycarbonate (PC) and polyethylene terephthalate (PET) with the thickness of 2.5 μm were used. The polymers used to synthesize these films are characterized by varying degrees of polarity, glass transition and softening temperatures as well as different morphology of the material structure.

The films were fixed in a special circular holder and placed between steel electrodes, one disc-shaped (40 mm in diameter) and the other sphere-shaped (6 mm in diameter). The surface of the electrodes was polished to a specular gloss. The space between the spherical electrode and the film was filled with capacitor oil to prevent edge and surface discharges. A high-voltage pulse of negative polarity with an amplitude of 2.5 kV and rise time of about 130 ns was applied once to the given film samples.

The electrode system used provided an electric field close to homogeneous in the region of film breakdown. Since breakdown of films occurred at the leading edge of the pulse, it is preferable to use the value of the breakdown strength F_{br} to evaluate their impulse strength, defining it as

$$F_{br} = U_{br}/d,$$

where U_{br} is the voltage in the sample at the time of breakdown, d is the thickness of the film.

U_{br} was measured with an ADS-2332 broadband storage oscilloscope and a high-voltage broadband divider with a cutoff frequency of 300 MHz, allowing to directly record the voltage variations in the sample. The time of film breakdown was detected in the waveforms by a sharp voltage drop and occurrence of oscillations. A typical waveform recorded at the time of breakdown at the leading edge of the pulse is shown in Fig. 1.

Electrical tests were carried out in the temperature range of 293–363 K. We should note that the chosen upper temperature limit was lower than the softening temperature of any of the polymers considered to prevent the spherical electrode from penetrating through the film with a consequent decrease in its breakdown voltage.

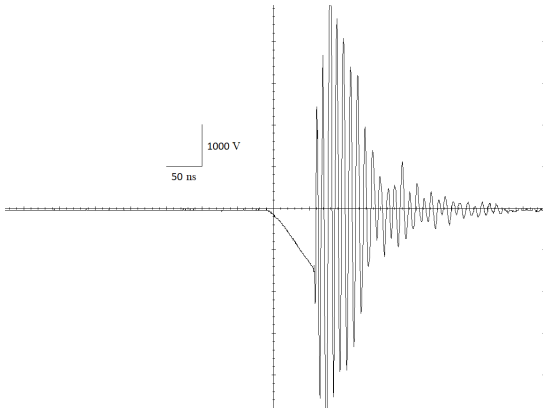

Fig. 1. Typical signal waveform at time instant of breakdown for 2.0 or 2.5 μm thick polymer film at pulse leading edge

Experimental results and discussion

The voltages U_{br} of polymer dielectric materials, including polymer films, are characterized by a significant scatter, therefore, breakdown tests should be repeated multiple times, and their results should be statistically processed. At least 50 samples of each polymer film were tested for breakdown at each temperature; the values of F_{br} found were used to calculate the empirical distribution function $f(F)$ determining the breakdown probability in polymer film when the voltage U (electrostatic field strength F) is reached. The values of the function $f(F)$ were calculated by the formula

$$f(F) = n/N, \tag{1}$$

where N is the number of tests performed, $N \geq 50$; n is the number of samples experiencing breakdown when the field strength F is reached.

It was established in [10–12] that the Weibull model of failure probability is applicable for statistical analysis of the results obtained for electrical breakdown under both DC and pulse voltages applied to polymer dielectrics. For this reason, a two-parameter Weibull distribution was used to approximate the empirical distribution function determined by relation (1), taking the form

$$f(F) = 1 - \exp[-(F/F_0)^m], \tag{2}$$

where m is the shape parameter, F_0 is the scale parameter.

The parameters m and F_0 are easily evaluated using the least squares method with expression (2) linearized.

Fig. 2 shows a typical form of the linearized functions $f(F)$ for polymer films studied for breakdown at different temperatures. The correlation coefficient of the approximating lines turned out to be higher than 0.97 in all cases, confirming the validity of the linear approximation used. The values of the parameters m and F_0 were calculated by the least squares method, and then used to calculate the first four moments μ_k of the distribution function [13]:

$$\mu_k = F_0^k \int_0^\infty x^{k/m} \exp(-x) dx. \tag{3}$$

This is the expected value of the Weibull distribution F_{br} (at $k = 1$), its variance σ (at $k = 2$), skewness μ_3 (at $k = 3$) and kurtosis μ_4 (at $k = 4$).

Fig. 3 shows the temperature dependences for the skewness and kurtosis of the function $f(F)$ of the studied polymer films. Evidently, $\mu_3 \neq 0$ for all temperatures, and the skewness values calculated at different temperatures are close; therefore, the functions $f(F)$ cannot be considered symmetric. However, the values of skewness μ_3 are small and predominantly positive, which indicates a slight positive skewness of the function $f(F)$.

The kurtosis μ_4 of this function varies from 2.5 to 3.5 in the studied temperature range, and does not depend on temperature. Note that the normal distribution should have zero asymmetry and a kurtosis equal to 3 [13]. The values of μ_3 and μ_4 that we calculated are close to these values, however, the breakdown voltage distributions that we obtained cannot be considered normal.

The temperature dependences of the most probable breakdown strength and the corresponding variances at each temperature are shown in Fig. 4 for all three types of polymer films. The values of F_{br} show a significant scatter for all polymers in the temperature range considered, showing a trend towards decrease with increasing temperature (at least for PP and PC films).

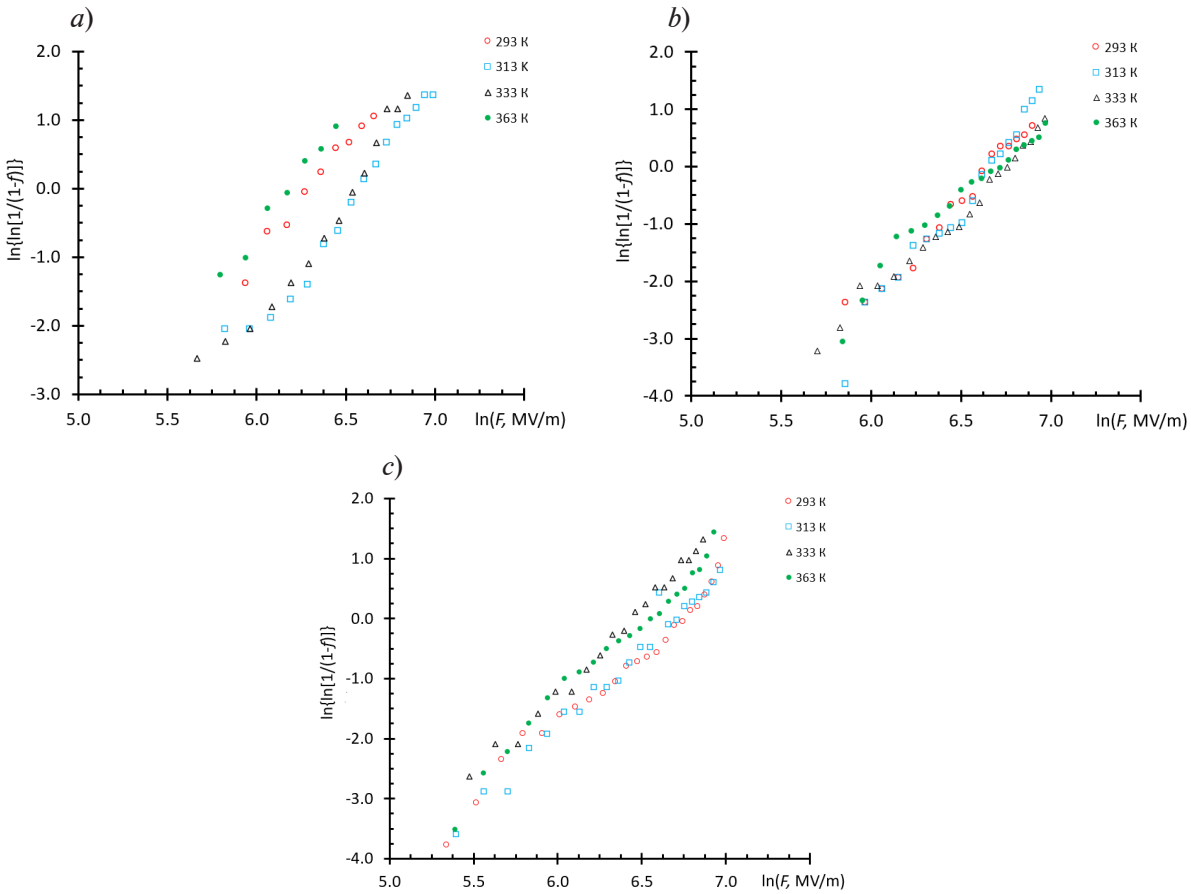

Fig. 2. Linearized Weibull distribution functions for PP (*a*), PET (*b*) and PC (*c*) films at different temperatures in the range of 293–363 K

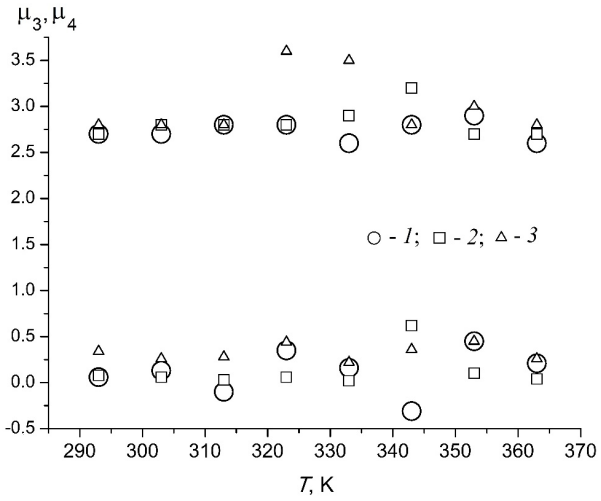

Fig. 3. Temperature dependences for skewness (μ_3, lower symbols) and kurtosis (μ_4, upper symbols) of breakdown strength distribution in PET films (symbols *1*), PP (*2*) and PC (*3*)

Student's *t*-test was applied to statistically validate the hypothesis of the decrease in F_{br} in polymer films under heating taking into account a small number of sample elements (determined by the number of temperature points for which the tests were carried out). The sample of F_{br} values for each polymer was divided into two groups with equal number of temperature points, low-temperature and high-temperature ones; next, Student's t-test (with a confidence interval of 90%) was used to compare the corresponding average values calculated for each of these groups [14]. The results obtained by this technique indicate that the hypothesis formulated about the effect of temperature on F_{br} can be accepted only for the PC film within the given confidence interval. The effect of temperature on F_{br} cannot be estimated as significant for PP and PET films.

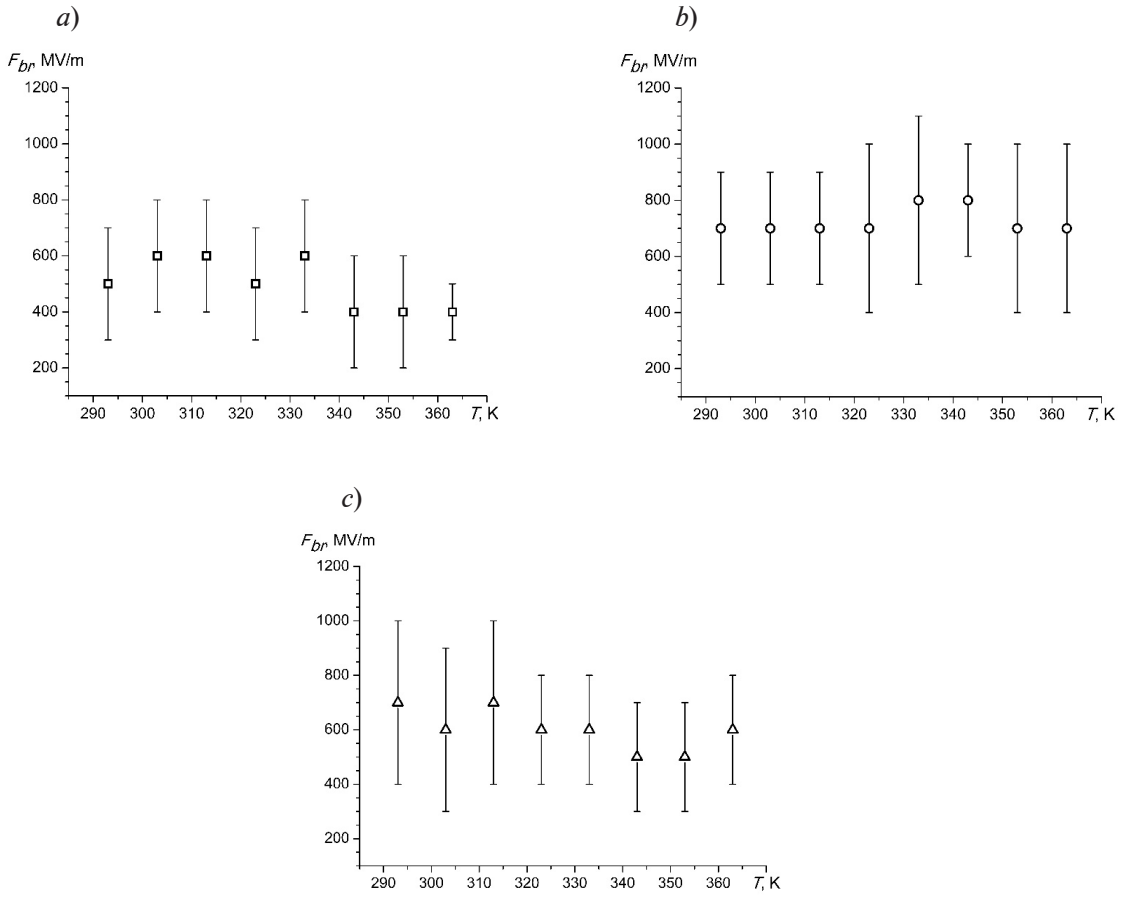

Fig. 4. Temperature dependences of most probable breakdown strength for PP (*a*), PET (*b*) and PC (*c*) films

Notably, neither the PP film nor the PC film changed their phase state during the heating process, since in the first case, the glass transition region is 253–263 K [15] and the PP film was in elastic state at a temperature above room temperature; on the contrary, in the second case, the glass transition temperature $T_g \approx 415$ K and the PC film was in a glassy state, even at the highest temperature of 363 K (used in our study). As for the PET film, its glass transition temperature is $T_g \approx 343$ K, i.e., T_g lies in the temperature range we considered but no noticeable change in F_{br} was observed at this point.

Additional data obtained during electrical tests of polymers in a DC field confirm that a sharp decrease in their breakdown strength occurs in the glass transition temperature region [16, 17]. Thus, it can be assumed that unfreezing of molecular mobility in polymers does not have a noticeable effect on impulse strength of polymer dielectric films, which is likely due to the short-term effect of the electric field on the polymer. Indeed, this time interval does not exceed 100 ns for film breakdown at the leading edge of the pulse.

It was found that the time interval for the final stage of electrical breakdown in polymer films, when fracture of the dielectric material occurs accompanied by the formation of a breakdown channel, is approximately 10^{-9}—10^{-8} s; the amplitude of breakdown current density reaches values of about 10^7 A/cm^2 [17, 18].

Note that the voltage oscillations in the sample, always detected in the waveforms after breakdown, can last for several hundred nanoseconds (see Fig. 1), but their occurrence is due to response of the measuring circuit to a short-term pulse of the breakdown current [18]. Evidently, finding out the reason behind such a rapid increase in current during breakdown remains one of the most important goals, which should allow to gain a better understanding of the picture of impulse breakdown in polymer dielectric films, as no consensus has been reached on this phenomenon.

Two possible physical mechanisms governing electrical breakdown in polymers are collisional ionization generating an electron avalanche [4, 9, 19, 20] and field-induced (tunneling) ionization of macromolecules [6—8]. We should note that the hypothesis of collisional ionization in polymers has faced much criticism in the recent years (see, for example, [6—8]). The cause for criticism is that the mean free path of electrons in polymer dielectrics does not exceed 1—3 nm, while electrons cannot gain the energy of 6—7 eV necessary for ionization of polymer macromolecules in a realistically achievable electric field.

On the other hand, the probability of large nanopores appearing in polymers increases with increasing temperature [21]. It is hypothesized in [22] that collisional ionization occurs precisely in such pores. The theory based on these hypotheticals [22] suggests a sharp dependence of breakdown strength on temperature, especially pronounced at temperatures above the glass transition temperature of the polymer. However, as mentioned above, the value of F_{br} is virtually independent of temperature in the case of impulse breakdown of polymer films.

As discussed in [6—8], the theory behind field-induced (tunneling) ionization of macromolecules can be used to explain the picture of impulse breakdown in polymers. According to this theory, electrons and positively charged molecular ions (holes) appear as a result of field-induced ionization of macromolecules. The explosive increase in their concentration is due to the Debye screening effect, manifesting when a certain critical charge density is reached. Debye screening leads to a decrease in the ionization energy of macromolecules and consequently to self-acceleration of the field-induced ionization process. The theory was expanded in [8], suggesting exposure to voltage pulses accelerates the field-induced ionization of macromolecules at the leading edge of the pulse. The reason for this effect is that electrons injected from the cathode can gain energy sufficient to excite molecules but not to ionize them in an electric field. Excited polymer molecules are ionized at lower field strengths than molecules in the ground state. Therefore, the critical concentration of electrons and holes is achieved in less time, which is the actual reason for breakdown at the leading edge of the pulse. According to the explanation given in [8], a slight decrease in the breakdown strength of polymer films under heating is associated with an increase in the free electron concentration both due to an increase in injected current density and due to an increased probability of thermally stimulated electron emission from traps.

Conclusion

This paper reports on the experimental study of the temperature effect on the breakdown strength of thin polymer dielectric films of polypropylene, polycarbonate and polyethylene terephthalate in the temperature range from 293 to 363 K. We found that the two-parameter Weibull distribution can be used for statistical processing of experimental results. The parameters of this distribution, determined at each temperature for each film, allow to calculate the most probable breakdown strength F_{br} and its variance σ. It was found that the value of F_{br} decreases slightly in polymers heated in the given range but is characterized by significant scatter that does not depend on temperature.

We established that the experimental data on the temperature effect on the impulse strength of polymers can be explained via the ionization mechanism of polymer breakdown, which is not associated with collisional ionization developing in the polymers.

REFERENCES

1. **Selle F., Schulte D.,** Electrical breakdown of EPR-insulated cables at DC and impulse voltages, IEEE Trans. Electr. Insul. 22 (2) (1987) 139–143.

2. **Watson D. B.,** Time lags in the impulse voltage breakdown of PMMA at high temperatures, J. Phys. D: Appl. Phys. 29 (12) (1996) 3129–3131.

3. **Krivko V. V., Lekht Yu. I.,** Vliyaniye temperatury na impulsnoye probivnoye napryazheniye nekotorykh tverdykh polimernykh dielektrikov [The temperature influence on the impulse voltage breakdown of some solid polymeric dielectrics], Bulletin of the Tomsk Polytechnic Institute. 157 (1970) 201–205 (in Russian).

4. **Chen Z., Ji S., Jia W., et al.,** Experimental study of the breakdown characteristics of polypropylene films under nanosecond voltage pulses, IEEE Trans. Plasma Sci. 46 (11) (2018) 4010–4016.

5. **Adam'yan Y. E., Krivosheev S. I., Magazinov S. G.,** Describing the impulse insulation strength of polymer dielectrics, Techn. Phys. Lett. 47 (3) (2021) 211–213.

6. **Zakrevski V. A., Pakhotin V. A., Sudar' N. T.,** Pulsed electric strength of polypropylene film, Techn. Phys. 62 (2) (2017) 276–281.

7. **Zakrevski V. A., Pakhotin V. A., Sudar' N. T.,** Effect of the voltage pulse front steepness on the electric strength of polymers, Techn. Phys. 63 (12) (2018) 1814–1819.

8. **Semenov S. E., Kapralova V. M., Pakhotin V. A., Sudar N. T.,** A mechanism of pulse breakdown evolution in polymeric films, Physics of the Solid State. 64 (8) (2022) 954–960.

9. **Zhao L., Su J. C., Liu C. L.,** Review of developments on polymers' breakdown characteristics and mechanisms on a nanosecond time scale, AIP Adv. 10 (3) (2020) 035206.

10. **Dissado L. A.,** Theoretical basis for the statistics of dielectric breakdown, J. Phys. D: Appl. Phys. 23 (12) (1990) 1582–1591.

11. **Wilson M. P., Given M. J., Timoshkin I. V.,** Impulse-driven surface breakdown data: A Weibull statistical analysis, IEEE Trans. Plasma Sci. 40 (10) (2012) 2449–2455.

12. **Zhao L., Su J. C., Zhang X. B., et al.,** Research on reliability and lifetime of solid insulation structures in pulsed power systems, IEEE Trans. Plasma Sci. 41 (1) (2012) 165–172.

13. **Kapur K. C., Lamberson L. R.,** Reliability in engineering design, Wiley, Hoboken, New Jersey, US, 1991.

14. **Berestneva O. G., Marukhina O. V., Shevelev G. E.,** Prikladnaya matematicheskaya statistika [Applied mathematical statistics], Tomsk Polytechnic Institute Publishing, Tomsk, 2012 (in Russian).

15. Propilena polimery; Entsiklopediya polimerov. V 3 tt. Glav. red. V. A. Kabanov. T. 3. [Encyclopedia of polymers. In 3 Vols, Chief Editor V. A. Kabanov, Vol. 3. Soviet Encyclopedia Publishing, Moscow, 1152 Columns (1977) 208–214 (in Russian).

16. **Sazhin B. I., Lobanov A. M., Romanovskaya O. S.,** Electrical properties of polymers, Third ed., Khimiya Publishing, Leningrad, 1986 (in Russian).

17. **Pakhotin V. A., Zakrevskii V. A., Sudar' N. T.,** Measurement of breakdown current in dielectric materials, Techn. Phys. 60 (8) (2015) 1144–1149.

18. **Pakhotin V. A., Sudar N. T.,** A technique for measuring the resistance of an electrical breakdown channel in thin dielectric films, Instruments and Experimental Techniques. 62 (3) (2019) 329–336.

19. **Zhao L.,** A formula to calculate solid dielectric breakdown strength based on a model of electron impact ionization and multiplication, AIP Adv. 10 (2) (2020) 025003.

20. **Zhao L.,** Theoretical calculation on formative time lag in polymer breakdown on a nanosecond time scale, IEEE Trans. Dielectr. Electr. Insul. 27 (4) (2020) 1051–1058.

21. **Tyutnev A. P., Sadovnichii D. N., Sayenko V. S., Pozhidayev E. D.,** The effect of molecular motions on the transport of excess charge carriers in polymers, Polymer Sci. 42A (1) (2000) 16–26 (in Russian).

22. **Artbauer J.,** Electric strength of polymers, J. Phys. D: Appl. Phys. 29 (2) (1996) 446–456.

СПИСОК ЛИТЕРАТУРЫ

1. **Selle F., Schulte D.** Electrical breakdown of EPR-insulated cables at DC and impulse voltages // IEEE Transactions on Electrical Insulation. 1987. Vol. 22. No. 2. Pp. 139–143.

2. **Watson D. B.** Time lags in the impulse voltage breakdown of PMMA at high temperatures // Journal of Physics D: Applied Physics. 1996. Vol. 29. No. 12. Pp. 3129–3131.

3. **Кривко В. В., Лехт Ю. И.** Влияние температуры на импульсное пробивное напряжение некоторых твердых полимерных диэлектриков // Известия Томского политехнического института. 1970. Т. 157. С. 201–205.

4. **Chen Z., Ji S., Jia W., Tang J., Guo F., Li J., Chen W.** Experimental study of the breakdown characteristics of polypropylene films under nanosecond voltage pulses // IEEE Transactions on Plasma Science. 2018. Vol. 46. No. 11. Pp. 4010–4016.

5. **Адамьян Ю. Э., Кривошеев С. И., Магазинов С. Г.** Особенности описания импульсной электрической прочности полимерных диэлектриков // Письма в Журнал технической физики. 2021. Т. 47. № 5. С. 48–50.

6. **Закревский В. А., Пахотин В. А., Сударь Н. Т.** Импульсная электрическая прочность пленки полипропилена // Журнал технической физики. 2017. Т. 87. № 2. С. 249–253.

7. **Закревский В. А., Пахотин В. А., Сударь Н. Т.** Влияние крутизны фронта импульса напряжения на электрическую прочность полимеров // Журнал технической физики. 2019. Т. 89. № 1. С. 120–125.

8. **Семенов С. Е., Капралова В. М., Пахотин В. А., Сударь Н. Т.** Механизм развития импульсного пробоя в полимерных пленках // Физика твердого тела. 2022. Т. 64. № 8. С. 948–954.

9. **Zhao L., Su J. C., Liu C. L.** Review of developments on polymers' breakdown characteristics and mechanisms on a nanosecond time scale // AIP (American Institute of Physics) Advances. 2020. Vol. 10. No. 3. P. 035206.

10. **Dissado L. A.** Theoretical basis for the statistics of dielectric breakdown // Journal of Physics D: Applied Physics. 1990. Vol. 23. No. 12. Pp. 1582–1591.

11. **Wilson M. P., Given M. J., Timoshkin I. V.** Impulse-driven surface breakdown data: A Weibull statistical analysis // IEEE Transactions on Plasma Science. 2012. Vol. 40. No. 10. Pp. 2449–2455.

12. **Zhao L., Su J. C., Zhang X. B., Pan Y. F., Wang L. M., Sun X., Li R.** Research on reliability and lifetime of solid insulation structures in pulsed power systems // IEEE Transactions on Plasma Science. 2012. Vol. 41. No. 1. Pp. 165–172.

13. **Капур К., Ламберсон Л.** Надежность и проектирование систем. Пер. с англ. М.: Мир. 1980. 604 с.

14. **Берестнева О. Г., Марухина О. В., Шевелев Г. Е.** Прикладная математическая статистика. Томск: Изд-во Томского политех. ун-та, 2012. 188 с.

15. Пропилена полимеры // Энциклопедия полимеров. В 3 тт. Гл. ред. В. А. Кабанов. Т. 3. Полиоксадиазолы – Я. М.: Советская энциклопедия, 1977. 1152 стб. С. 208–214.

16. **Сажин Б. И., Лобанов А. М., Романовская О. С.** Электрические свойства полимеров. Под ред. Б. И. Сажина. 3-е изд. Ленинград: Химия, 1986. 224 с.

17. **Пахотин В. А., Закревский В. А., Сударь Н. Т.** Об измерении тока электрического пробоя диэлектриков // Журнал технической физики. 2015. Т. 85. № 8. С. 40–45.

18. **Пахотин В. А., Сударь Н. Т.** Методика измерения сопротивления канала электрического пробоя тонких диэлектрических пленок // Приборы и техника эксперимента. 2019. № 3. С. 30–37.

19. **Zhao L.** A formula to calculate solid dielectric breakdown strength based on a model of electron impact ionization and multiplication // AIP (American Institute of Physics) Advances. 2020. Vol. 10. No. 2. P. 025003.

20. **Zhao L.** Theoretical calculation on formative time lag in polymer breakdown on a nanosecond time scale // IEEE Transactions on Dielectrics and Electrical Insulation. 2020. Vol. 27. No. 4. Pp. 1051–1058.

21. **Тютнев А. П., Садовничий Д. Н., Саенко В. С., Пожидаев Е. Д.** Молекулярные движения и их роль в переносе избыточных зарядов в полимерах // Высокомолекулярные соединения. Серия А. 2000. Т. 42. № 1. С. 16–26.

22. **Artbauer J.** Electric strength of polymers // Journal of Physics D: Applied Physics. 1996. Vol. 29. No. 2. Pp. 446–456.

THE AUTHOR

SEMENOV Sergey E.
Peter the Great St. Petersburg Polytechnic University
Ioffe Institute, RAS
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
moritohayama96@gmail.com
ORCID: 0009-0000-5779-9966

СВЕДЕНИЯ ОБ АВТОРЕ

СЕМЕНОВ Сергей Евгеньевич – *аспирант Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого, младший научный сотрудник Физико-технического института* им. А.Ф. Иоффе РАН.
195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
moritohayama96@gmail.com
ORCID: 0009-0000-5779-9966

Received 12.01.2024. Approved after reviewing 22.01.2024. Accepted 01.02.2024.
Статья поступила в редакцию 12.01.2024. Одобрена после рецензирования 22.01.2024. Принята 01.02.2024.

Journal

ST. PETERSBURG STATE POLYTECHNICAL UNIVERSITY JOURNAL: PHYSICS AND MATHEMATICS

Vol. 17, No. 2, 2024

Founder and publisher: Peter the Great St. Petersburg Polytechnic University

The journal is registered with the Federal Service for Supervision of Communications, Information Technology and Mass Media (Roskomnadzor).
Certificate ПИ ФС77-51457 issued 19.10.2012.

Phone 8 (812) 294-22-85

Website https://physmath.spbstu.ru/

E-mail: physics@spbstu.ru

Typesetting by A.S. Kolgatina

Published 30.06.2024. Format 60x84/8. Digital print.
Printer's sheets Print circulation 1000. Order ID