

## ФИЗИКА

УДК 535.4; 535.317

### BALLISTIC AUTO-CORRELATION INTERFEROMETRY

P. Vacas-Jacques, V. Ryabukho<sup>\*,\*\*</sup>, M. Strojnik, V. Tuchin<sup>\*,\*\*</sup>, G. Paez

Centro de Investigaciones en Optica, 37000, Leon, Guanajuato, Mexico

E-mail: pvj@cio.mx

\* Saratov State University, Russia

E-mail: rvp@sgu.ru

\*\* Institute of Precision Mechanics and Control of the Russian Academy of Sciences, Saratov, Russia

We propose the implementation of a ballistic auto-correlation interferometer (BAI) for the assessment of biological tissues. We develop the theory of the interferometric technique, and demonstrate that a method to isolate ballistic from forward-scattered contributions is of central importance. We, therefore, propose theoretically and experimentally a non-linear grating-based angular filter. A monochromatic source incident on a ruled grating, positioned at grazing diffraction, is followed by a narrow slit to conform the proposed system. We validate the angular amplification experimentally, with values ranging on the order of 10–20X. In addition, similar values of transversal beam size reduction, provide an efficient ~100X filtering scheme. As a preliminary study for the implementation of the BAI, we compare two transillumination schemes, which include and exclude the grating-based angular filter. Our preliminary results are encouraging, and indicate that the filter effectively isolates the information needed in the BAI.

**Key words:** interferometry, ballistic transillumination, angular filter, grazing diffraction, tissue imaging, tissue characterization.

#### Баллистическая автокорреляционная интерферометрия

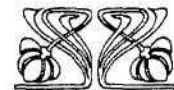
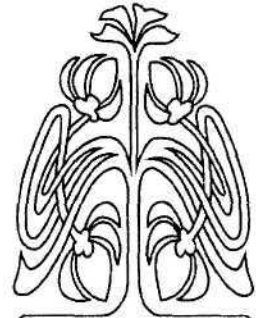
П. Вакас-Жакс, В. Рябухо, М. Стройник, В. Тучин, Г. Паез

Рассматривается применение баллистической автокорреляционной интерферометрии (БАИ) для контроля параметров биологических тканей. Представлена теория интерференционного эксперимента и показано, что автокорреляционный метод интерферометрии позволяет изолировать и выделить в рассеянном вперед излучении нерассеянное, баллистическое, излучение. Для реализации метода предложено использовать нелинейный дифракционный угловой фильтр, действие которого основано на пространственной фильтрации света с помощью узкой щели света в скользющем дифракционном порядке. Реализованное с помощью такого фильтра угловое увеличение составило 10–20X. Такое же увеличение дает поперечное масштабирование дифрагированного пучка света, что в целом обеспечивает эффективную ~100X схему угловой фильтрации рассеянного излучения. Выполнено сравнение двух схем БАИ с угловым фильтром и без его использования. Экспериментальные результаты подтверждают теоретические положения и показывают эффективное действие фильтра для реализации метода БАИ.

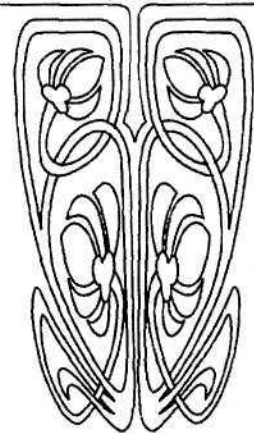
**Ключевые слова:** интерферометрия, баллистическое просвечивание, угловой фильтр, дифракция, имиджинг биологических тканей, диагностика ткани.

#### 1. Introduction

Transmission of light in a turbid medium, such as tissue, is governed by the radiative transport equation [1–4]. The attenuation characteristics of the media are required to describe the propagation of radiation [5]. More specifically, the necessary parameters are the absorption coefficient  $\mu_a$  [ $\text{mm}^{-1}$ ], the scattering coefficient  $\mu_s$  [ $\text{mm}^{-1}$ ], and the anisotropy factor  $g$ . The first two coefficients describe the probability of absorption and scattering of light per unit length. Meanwhile, the third parameter determines the average direction of propagation after interaction. In addition, the anisotropy of propa-



НАУЧНЫЙ  
ОТДЕЛ





gation after interaction. In addition, the anisotropy factor is related to a specific phase function, which accounts for the pattern of scattering distinctive of the turbid medium. Biological tissues commonly exhibit phase functions with marked forward-directed patterns [6, 7]. Therefore, incident and scattered propagation directions are very similar.

One of the challenges of implementing a transillumination technique with ballistic photons resides in the rapid attenuation of such ra-

diation. A perhaps more complicated endeavor however, is the separation of ballistic photons from those that have undergone scattering. The latter assertion is further validated by acknowledging that the anisotropy factor is usually greater than 0.9 for biological tissues [7]. In Fig. 1, we give an example of the intermingled presence of ballistic and scattered photons after transillumination. The sample is 10 mm thick;  $\mu_a = 0$  [mm<sup>-1</sup>],  $\mu_s = 0.591$  [mm<sup>-1</sup>], and  $g = 0.93$ .

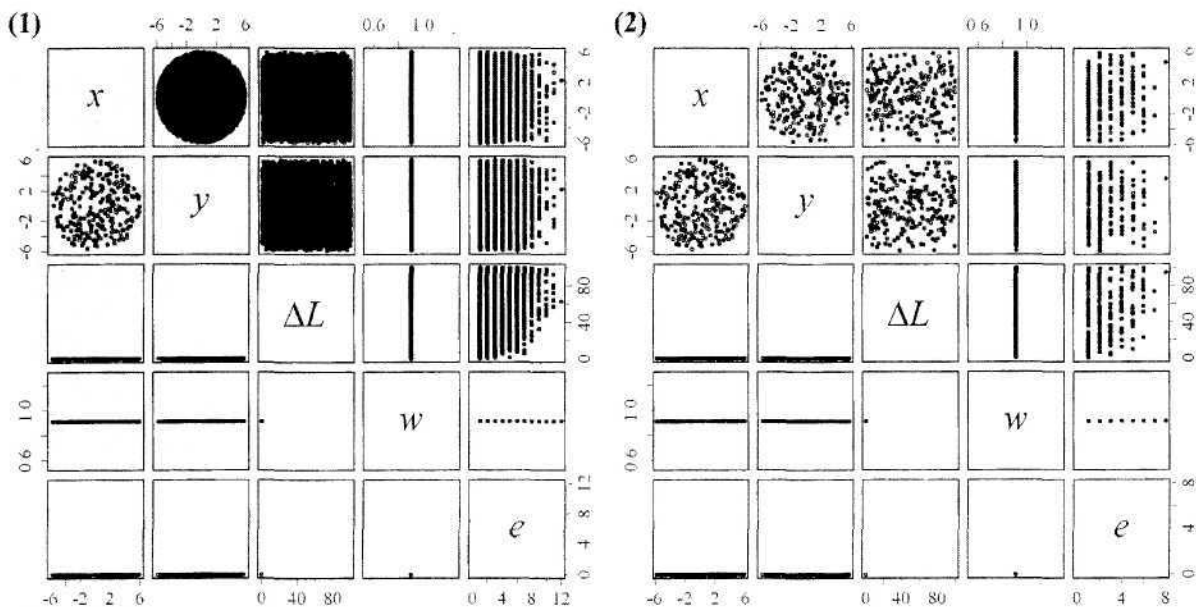


Fig. 1. Ballistic and scattered contributions are commingled after transillumination: immediate (1) and distant (2) observation

In Fig. 1, data below and above the main diagonal correspond to that of ballistic and scattered photons. Each row-column intersection is a scatter plot of the associated variables in the diagonal.  $x$  and  $y$  are transversal coordinates (in [mm]),  $\Delta L$  represents optical path-length difference (in [ $\mu$ m]),  $w$  is weight, and  $e$  stands for the number of scattering events. As seen from Fig. 1, distinction of ballistic from scattered photons is challenging. Despite the observation position, near (1) or far (2), both contributions lie close together. Such behavior of radiation transport is relevant in the ballistic auto-correlation interferometer (BAI). In Fig. 2, we depict a schematic of the BAI.

In the following section, we present the theoretical foundation of the BAI. We emphasize the importance of adequately selecting the

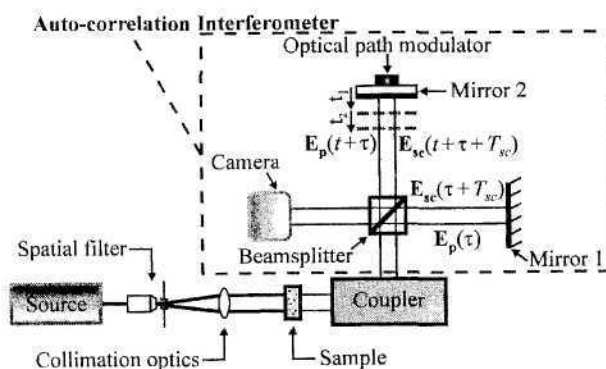


Fig. 2. In the BAI a coupler is necessary to isolate ballistic from forward-scattered radiation

information of interest. For this purpose, we suggest the implementation of a non-linear grating-based angular coupler. In the third part of this work, we describe experimental setup and



methodology. In section four, we then assess the performance of the angular coupler. In addition, we compare two transillumination schemes, which include and exclude the grating-based system. We devote the fifth section to discuss possible improvements. Finally, we summarize the results of this work and present plausible future research directions.

## 2. Theory

As pointed out in the introduction, the radiation impinging on the sample suffers diverse interactions in the turbid medium. As a result of such interactions, the radiation is transmitted, reflected, or absorbed. In essence, we may speak of two contributions that are transmitted: forward-scattered and ballistic. The transmitted components are especially relevant for the implementation of a ballistic auto-correlation interferometer. The coherent superposition of transmitted beams is employed in the BAI to enable characterization and imaging of turbid media. Electrical fields are associated with incident  $\mathbf{E}_i$ , forward-scattered  $\mathbf{E}_{sc}$ , and ballistic or pass-through  $\mathbf{E}_p$  contributions. At the superposition plane of the BAI, see Fig. 2, ballistic and forward-scattered components are observed:

$$P(t, T_{sc}) \propto \left\langle \left\{ \left[ \mathbf{E}_p(\tau) + \mathbf{E}_{sc}(\tau + T_{sc}) \right] + \left[ \mathbf{E}_p(t + \tau) + \mathbf{E}_{sc}(t + \tau + T_{sc}) \right] \right\} \times \left\{ \left[ \mathbf{E}_p(\tau) + \mathbf{E}_{sc}(\tau + T_{sc}) \right] + \left[ \mathbf{E}_p(t + \tau) + \mathbf{E}_{sc}(t + \tau + T_{sc}) \right] \right\}^* \right\rangle. \quad (1)$$

Here,  $P$  stands for detected power.  $\tau$  is the reference time for partial coherence modeling.  $t$  represents the modulation time, and  $T_{sc}$  describes

$$P(t, T_{sc}) = \Gamma_{i,i}(0) \left\{ (1 - k_{sp1})(1 - k_{sp2}) \times \left( |\gamma_m|^2 + |\gamma_{sc}|^2 + 2|\gamma_m||\gamma_{sc}| |\gamma_{i,i}(T_{sc})| \cos[\Phi_{i,i}(T_{sc})] \right) + k_{sp1}k_{sp2} \left( |\gamma_m|^2 + |\gamma_{sc}|^2 + 2|\gamma_m||\gamma_{sc}| |\gamma_{i,i}(T_{sc})| \cos[\Phi_{i,i}(T_{sc})] \right) + 2\sqrt{k_{sp1}}\sqrt{1 - k_{sp1}}\sqrt{k_{sp2}}\sqrt{1 - k_{sp2}} \times \left( \left( |\gamma_m|^2 + |\gamma_{sc}|^2 \right) |\gamma_{i,i}(t)| \cos[\Phi_{i,i}(t)] + |\gamma_m||\gamma_{sc}| |\gamma_{i,i}(T_{sc} - t)| \cos[\Phi_{i,i}(T_{sc} - t)] + |\gamma_m||\gamma_{sc}| |\gamma_{i,i}(T_{sc} + t)| \cos[\Phi_{i,i}(T_{sc} + t)] \right) \right\}. \quad (3)$$

Here,  $\Gamma_{i,i}(0)$  is the total collected power upon proving a small area.  $k_{sp1}$  and  $k_{sp2}$  denote splitting coefficients. In the Michelson configuration,  $k_{sp2} = 1 - k_{sp1}$ . Finally,  $|\gamma_{i,i}(\tau)|$  and  $\Phi_{i,i}(\tau)$  designate amplitude and phase of the complex degree of self-coherence.

the random time that scattered photons take to traverse the sample. The chevrons denote time average and an asterisk demarcates complex conjugation. Throughout this work, we use bold letters for complex-valued quantities.

In Eq. (1), there are no reference contributions. This is one principal characteristic of the auto-correlation interferometer [8, 9]. In addition, we ascertain that forward-scattered and pass-through contributions are commingled. We may further simplify the expression of interference by relating the ballistic field to the incident one:

$$\mathbf{E}_p = R_g \exp\left[-i\frac{2\pi}{\lambda}2n_g d\right] \exp\left[-\frac{1}{2}\int_0^D \mu_t(z)dz\right] \times \exp\left[-i\frac{2\pi}{\lambda}\int_0^D n_s(z)dz\right] \mathbf{E}_i. \quad (2)$$

Here, the Fresnel reflection coefficient is  $P_g$  and we denote the wavelength as  $\lambda$ . The refractive index and thickness of the container, for in-vitro studies, are  $n_g$  and  $d$ . In addition, a sample of thickness  $D$ , attenuation coefficient  $\mu_t = \mu_a + \mu_s$ , and refractive index  $n_s$  is considered. In previous publications, we presented more succinctly the expression in Eq. (2) as  $\mathbf{E}_p = \gamma_m \mathbf{E}_i$  [10, 11]. Nonetheless, the complete equation serves to ascertain the rapid decay of ballistic radiation.

In contrast to the behavior of ballistic radiation, scattered contributions are random. Thus a simple expression between scattered and incident fields is not at hand. However, the fields are related to each other. Previously, we have employed the notation  $\mathbf{E}_{sc} = \gamma_{sc} \mathbf{E}_i$  for such purpose [10, 11]. By relating ballistic and scattered contributions to the incident field, we obtain the partial coherence model of the BAI:

From Eq. (3), we predict commingled detection of forward-scattered and ballistic photons. Therefore, we ascertain the necessity of adequately coupling the information of interest. In addition, from Fig. 1 we may identify the major competing process in the BAI. For this pur-



pose, we analyze the weight to number of events plot ( $w$  vs  $e$ ) for scattered photons. Such information is found in the fourth row and fifth column intersection. For simplicity, we examine the data of scenario (2). From the plot we establish that scattered photons with few events are the main competing process. The latter assertion is valid, because such photons maintain a propagation direction similar to that of ballistic radiation. Furthermore, the weight of such scattered radiation is significant, as opposed to that of highly diffuse contributions.

Particularly, single scattered radiation represents a diffuse contribution that is difficult to overcome. Single scattered photons, especially in highly forward-directed samples, do not suffer an important phase delay. Thus ballistic techniques which employ coherence gating, or ultra-fast imaging, encounter challenges while dealing with such radiation [10, 11]. However, single scattered photons indisputably modify their propagation direction. Hence, a technique which exploits such angular differences could be optimal to feed the BAI. For this purpose, we suggest the use of a ruled diffraction grating as coupling device. The peculiarity of our proposal resides in the regime of operation of the optical element. The grating is positioned such that one of the first diffraction orders  $m = +1$  occurs at a grazing angle, as shown next.

The behavior depicted in Fig.3 is supported by the theory of diffraction gratings. In fact, functioning of the coupler is based on conical diffraction theory. For in-plane geometries, diffracted and incident beams are related by the grating equation [12]:

$$\sin \beta = \sin \alpha + Gm \lambda . \quad (4)$$

Consider in Fig.3 that the incidence angle  $\alpha$  is  $13.8^\circ$  and that He-Ne  $\lambda = 632.8$  [nm] light is incident on the grating, which has a groove frequency  $G$  of 1200 [1/mm]. Then, the first diffraction order ( $m = +1$ ) occurs roughly at  $\beta = 86^\circ$ . By analyzing figures 1, 2, and 3, we note that the grazing diffraction grating might be an adequate coupler for the BAI. The latter assertion is valid, provided that the diffractive element is sensitive to minute variations in incidence angles. Mathematically, this is assessed by finding the derivative of Eq.(4):

$$\Delta \beta = \left( \frac{\cos \alpha}{\cos \beta} \right) \Delta \alpha . \quad (5)$$

The importance of Eq. (5) is that angular variations in the diffracted beam are amplified nonlinearly by  $(\cos \alpha / \cos \beta)$ . This assertion is especially evident for near grazing diffraction geometries,  $\beta \rightarrow \pi/2$  [rad]. In fact, for the parameters of the coupler mentioned before, the amplification is 15X. In Fig. 4, we present the behavior of scattered and ballistic radiation, before and after interaction with such coupler.

By inspection of the setup in Fig. 3, we observe a reduction of the transversal beam size. The decrease in beam size is also given by  $(\cos \alpha / \cos \beta)$ . For the case under consideration, such reduction is  $\sim 15X$ . Both angular amplification and decrease in beam width affect scattered photons. Therefore, in the BAI, the use of a narrow slit in conjunction with a grazing diffractive element may provide a filtering scheme of  $\sim 225X$ ;  $(\cos \alpha / \cos \beta)^2$ . This behavior, in turn, implies that ballistic photons may be exclusively coupled to the interferometer. Therefore, Eq. (3) simplifies accordingly.

$$P(t) = \left| \gamma_m \right|^2 \Gamma_{i,i} (0) \left\{ (1 - k_{sp1}) (1 - k_{sp2}) + k_{sp1} k_{sp2} + 2 \sqrt{k_{sp1}} \sqrt{1 - k_{sp1}} \sqrt{k_{sp2}} \sqrt{1 - k_{sp2}} \times \right. \\ \left. \times \left| \gamma_{ii}(t) \right| \cos \left[ \Phi_{i,i}(t) \right] \right\} . \quad (6)$$

By means of the expression in Eq. (6), we may identify several regimes of operation for the BAI. First, tissue characterization may be rea-

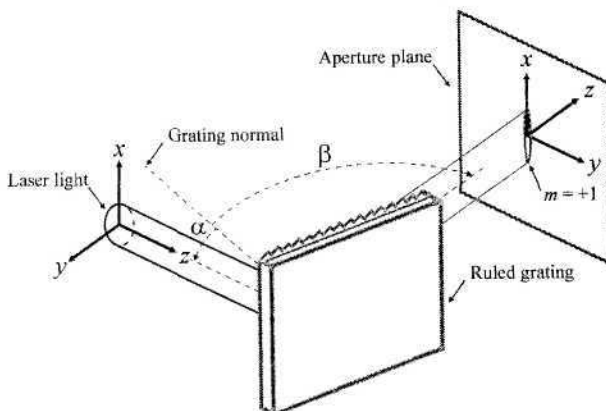


Fig. 3. A grazing diffraction element followed by a narrow slit conform the angular coupler

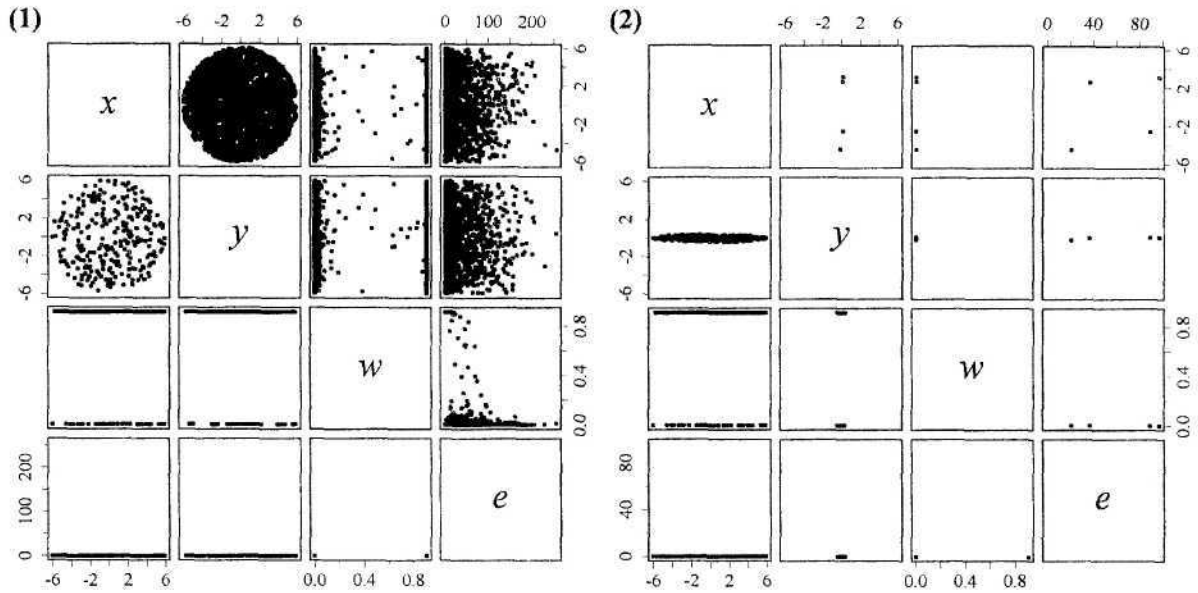


Fig. 4. Exclusion of the grazing grating (1) implies commingled presence of ballistic and scattered radiation, whereas inclusion of the 15X coupler (2) provides a means to overcome noise

lized by employing phase-locked techniques. In order to perform the latter, an accurate modulation reference is necessary. The information recovered by such method is proportional to  $|y_m|^2$ . In addition, a similar approach may yield the phase of the object under study. For this case, however, instead of phase-locked detection schemes, a phase-stepping approach is required. The second regime of operation corresponds to that of tissue imaging. From Eq. (6), we ascertain that pixel maxima are related to the integrated attenuation of the sample. In addition, another important aspect related to the application of tissue imaging, concerns the reduction of beam extent implicit in the configuration. Such beam reduction must be compensated accordingly, in order for the tissue imaging regime to be applicable. The prime issue to address, nonetheless, is the feasibility of implementing the grating-based angular filter.

### 3. Materials and Methods

**3.1. Experimental setup.** From theory, we expect a non-linear relation between diffracted and incident beams. The latter assertion is particularly relevant for the purposes of implementing a BAI. We are, therefore, interested in finding the rate of change of diffraction angle as a function of incidence angle. As a consequence, an important element of the experiment is the

rotating scheme employed to vary the incidence angle,  $\alpha \pm \Delta\alpha$ . We utilize a  $360^\circ$  rotation stage, which provides a  $15^\circ$  interval with 1 arc min resolution (Precision rotary stage 7R129, Standa Optics, Lithuania).

In addition, as the coupler approaches the grazing condition, we expect the diffracted angle to increase considerably. Thus, in the setup, another important component is the optical system needed to capture such variations. The optical system we employ exhibits a variable f-number. Furthermore, the diaphragm enables accurate constriction of the system to the paraxial regime. This regime is of importance, because accurate system alignment is facilitated. Accurate alignment is in turn required to calculate, by means of geometrical optics, the angular deviation of the diffracted beam. In Fig. 5, we depict a schematic of the experimental setup.

From Fig. 5, we recognize several important features related to the experiment. First, the mechanics of the rotating element play an important role when fine movements are in order. A mechanical malfunctioning, for instance a non-continuous movement, will manifest as a deviation from the expected performance. Second, the f-number of the optical system determines the incidence angle range that may be assessed. Thus if a considerable range is in play, it must be subdivided accordingly. Third, in order to determine the angular change after dif-

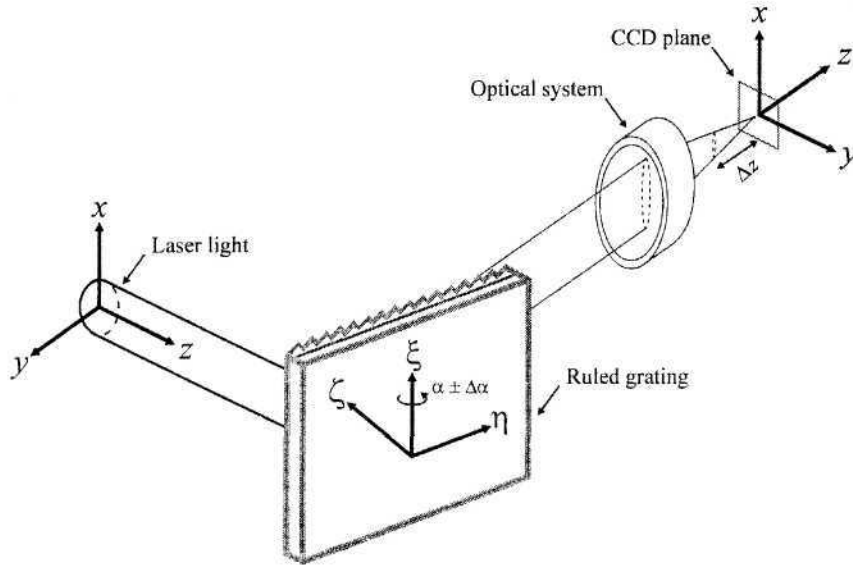


Fig. 5. Schematic representation of the system employed to validate the non-linear behavior of the grating-based angular coupler

fraction, we require the magnification of the optical system. In Fig. 6, we show a picture of the arrangement as implemented in our laboratory.

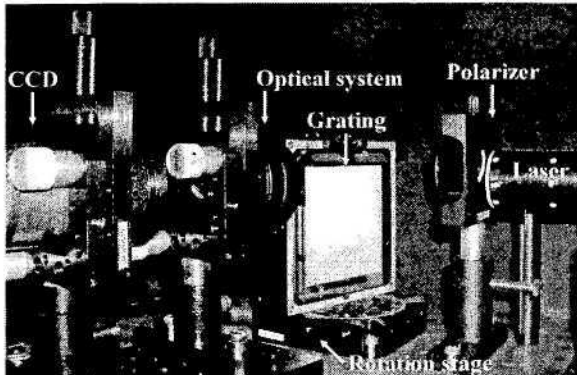


Fig. 6. Experimental setup utilized to validate the non-linear behavior of the proposed coupler

**3.2. Methodology.** Due to the fact that the aperture of the optical system is finite, we subdivide the range of incidence angles of interest. We select a determined angular interval and align the optical system, including the CCD. We perform the latter by working on the paraxial regime of the optical system. Once the alignment is achieved, we verify the latter by displacing the CCD (along the  $z$ -axis in Fig. 5), which is mounted on a translational stage. This procedure is important because the diffracted angle is calculated by using geometrical optics and trigonometry. The translational stage provides one of

the parameters needed in the calculations: the transversal reference position of the CCD. Meanwhile, we obtain the other information from CCD measurements.

From Fig. 5, we observe that all diffracted beams should converge in the image plane. In terms of geometrical optics, this is readily explained because we are tracing exclusively one ray. The latter, as shown in Fig. 5, always leaves the object plane (i.e., the diffraction grating) from the same place. After determining the location of convergence, we move the CCD a determined amount,  $\Delta z = 3$  [mm], along the optical axis. Thereafter, we calculate the transversal displacement of the beam in the CCD plane. In Fig. 7, we show a selected set of measurements from one of the experiments.

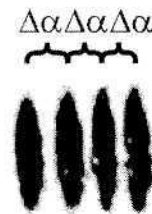


Fig. 7. Transversal displacement of diffracted beam as incidence angle varies  $\Delta\alpha = 7$  [arc min]

Depending on the angular range under consideration, we perform a sequence of measurements with an angular interval of 1 arc min.



Thus  $\Delta\alpha = 1/60\epsilon$  in Eq. (5), which implies that the derivative may be calculated accurately. Nonetheless, in the experiment, mechanical imperfections restrict the minimum feasible range. In our case, we space the measurements from five to eleven arc min. Worth of mention is that each sequence of measurements is repeated 10 times to account for statistical uncertainties.

From the alignment procedure, we determine the reference transversal position and the movement induced along the optical axis of the CCD. In addition, from the sequential variation of the incidence angle, we calculate the corresponding transversal displacement of the diffracted beam (see Fig. 7). Thereafter, we compute the values of diffraction angles by employing trigonometry and the magnification factor of the optical system. Finally, from successive measurements, we calculate the response of the system.

#### 4. Results

##### 4.1. Non-linear grating-based angular filter.

In the experiment, the exact value of the incidence angle  $\alpha$  is unknown. The angular increment  $\Delta\alpha$ , in contrast, is under control of the experimenter. From the measurements, we determine the relative increment in diffracted angle  $\Delta\beta$ . This conjoint information yields the sought rate of change,  $\Delta\beta/\Delta\alpha$ . In addition, as shown in Eqs. (4) and (5), the diffractive configuration and the incidence angle may also be employed to calculate the rate of change. Since the incidence angle is not known, we compute the theoretical behavior from the rate of change obtained experimentally. In the theoretical calculations, we employ the greatest value of  $\Delta\beta/\Delta\alpha$  as reference. From such value of the rate of change, we compute the corresponding incidence angle, and the theoretical curve for the configuration. In Fig. 8, we compare the results of the experiment with theory for two amplification schemes: a) 8–14X and b) 10–22X.

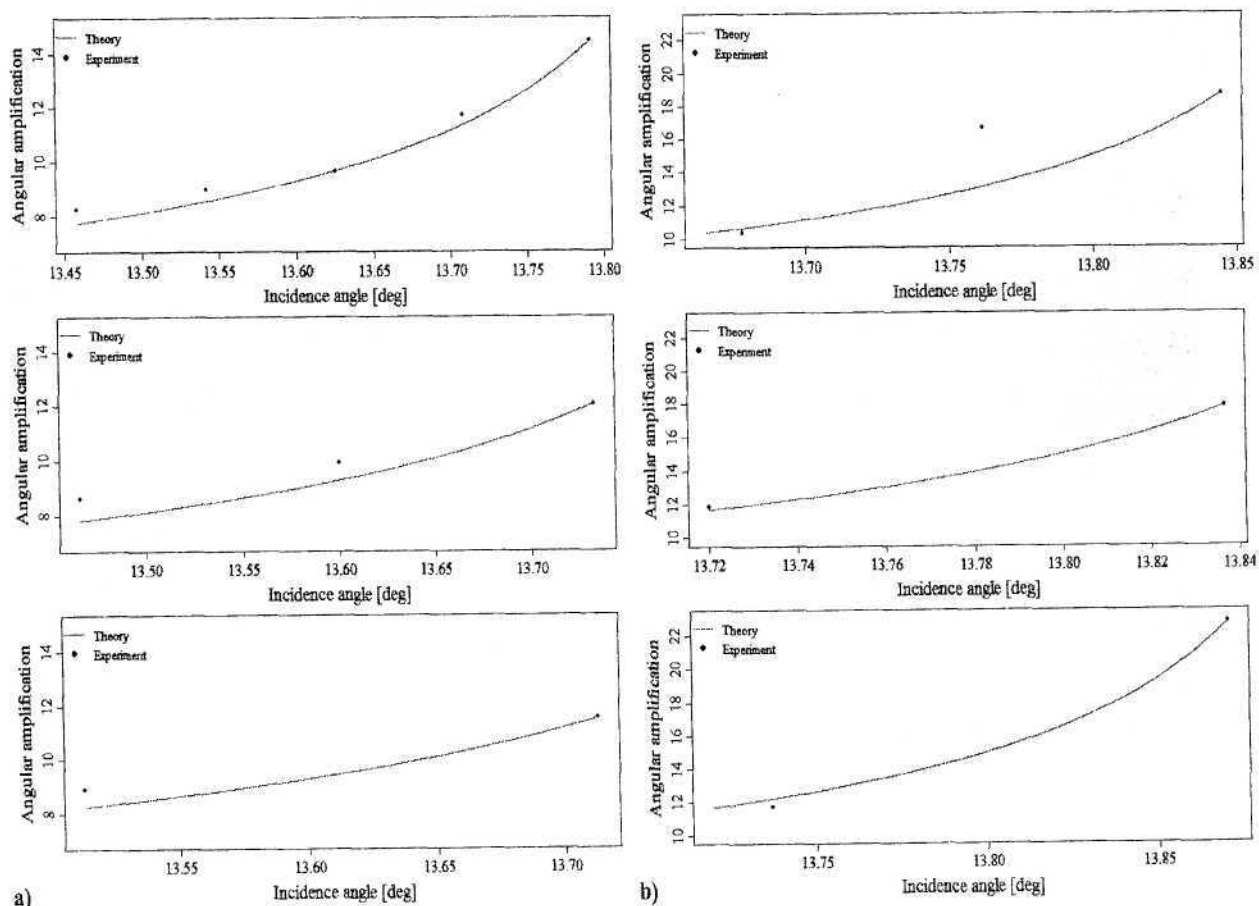


Fig. 8. Comparison of experiment with theory for two configurations: a) 8–14X and b) 10–22X



The angular amplification depicted in Fig. 8 is the rate of change of Eq. (5). From the figure, we ascertain agreement between theory and experiment. For the case of lower amplification, theoretical and experimental values exhibit good correspondence. Certain experimental values in the greater amplification scheme tend to oscillate more. These discrepancies are readily justified by acknowledging the non-linear behavior of the system.

In Fig. 8, the non-linear behavior of the system is evident. In addition, a relevant feature of the design of the experiment may be recognized from the same representation. Namely, marked non-linear response is confined to a few arc minutes. For example, if  $\alpha = 13.8^\circ$  then  $\Delta\beta/\Delta\alpha \sim 15$ . However, if we change the incidence angle by roughly 7.5 arc min, the grazing condition is exceeded. Due to the specifications of our equipment, the highly non-linear behavior of the filter was not assessed. Recently, in the field of spaceborne angular sensors, such regime has been explored [13, 14]. Nonetheless, from Fig. 8, we may conclude that amplification values ranging from 10–20X are feasible with 1 arc min resolution mounts.

**4.2. Transillumination: traditional versus grating coupler schemes.** We now address a topic that is closely related to the BAI. In order to

assess the applicability of the grating coupler, we perform two transillumination experiments. In the first scenario, we perform traditional transillumination. In Fig. 2, this experiment entails the replacement of the coupling device by a CCD or a camera. In the second setting, we implement the grating-based scheme. The coupler in Fig. 2 is replaced by the diffractive element at grazing regime. In addition, we employ a set of cylindrical lenses (in telescope arrangement) and a CCD to recover sample information [15].

The sample is composed of 1.5  $\mu\text{m}$  polystyrene latex spheres suspended in deionized water. The choice of particle size is related to the anisotropy factor, which is a central parameter in the BAI. We employ Mie theory to calculate the attenuation behavior of the sample [16]. Specifically, the anisotropy factor equals 0.935; a value characteristic of forward-scattering media. We employ a rather transparent sample, which is similar to the scenario encountered in dental enamel tissues [17]. The sample is held in a cuvette. Furthermore, a line label of such container is employed as the reference object. Images of reference object, traditional transillumination (employing only a CCD and a camera), and grating-based schemes are shown in Fig. 9.

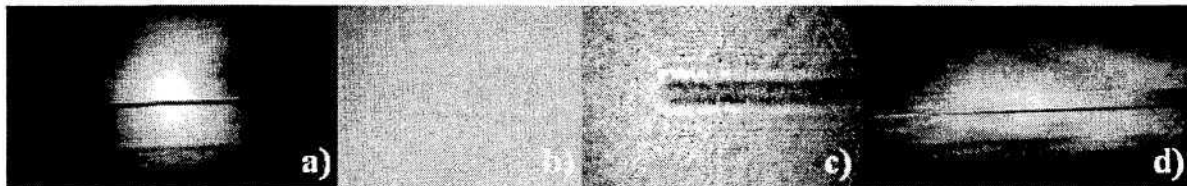


Fig. 9. Transillumination results: a) Reference, b) Traditional only CCD, c) Traditional with camera, d) Grating-based scheme

We concentrate on the salient features of the images presented in Fig. 9. First, in the traditional transillumination schemes, Figs. 9 b) and c), we observe a significant amount of forward-scattered contributions. Second, speckle patterns are also recognizable from these representations. These patterns may be more easily identified in a sequence of frames. Third, the exposition time necessary for acquiring the figures in d) is more significant than the corresponding measure for traditional transillumination. More specifically, the exposition time  $t_{\text{exp}}$  (VS-CTT-259-2001, Videoscan, Russia) is calculated from the rela-

tion in Eq. (7). In Table, we present the corresponding measures for the images in Fig. 9.

$$t_{\text{exp}} = 20000 \times TVSS + 64 \times THSS + 45.2 \text{ } [\mu\text{s}]. \quad (7)$$

Exposition times for the transillumination results of Fig. 9

Arrangement	TVSS	THSS	$t_{\text{exp}}$ [ $\mu\text{s}$ ]
Traditional only CCD*	0	33	2160
Traditional with camera*	0	9	620
Grating-based scheme	7	1	140110

\* THSS values for traditional modalities are approximate.





Due to the presence of the diffractive element, the grating-based configuration exhibits an implicit reduction of power at the CCD. Despite this fact, the exposition times shown in Table, for the grating-based scheme, are orders of magnitude greater than those of traditional transillumination. Such behavior is characteristic of the rapid decay of ballistic radiation. Therefore, these results suggest that the proposed method may be suited for coupling selectively the radiation required in the BAI.

### 5. Discussion

The device we have introduced in this work may be implemented readily for applications involving tissue characterization. In our group, we are also interested in performing tissue imaging. Furthermore, we would like to complement the

technique with interferometric methods. For imaging applications, we foresee challenges due to the marked asymmetry of the spatial distribution, see Fig. 4. In order to overcome such asymmetry, we may introduce a second diffractive element.

Taking Fig. 3 as reference, we should position the additional grating in the aperture plane. The grooves need to be parallel to the  $y$  axis, and by rotating the element we should reach the grazing diffraction condition. An advantage of this configuration is that the angular amplification occurs in a transversal direction. In addition, if the characteristics of the gratings match, we anticipate that the spatial asymmetry will be removed. This behavior implies that circular diaphragms may be used to block scattered light. In Fig. 10, we depict the theoretical response of the improved coupler.

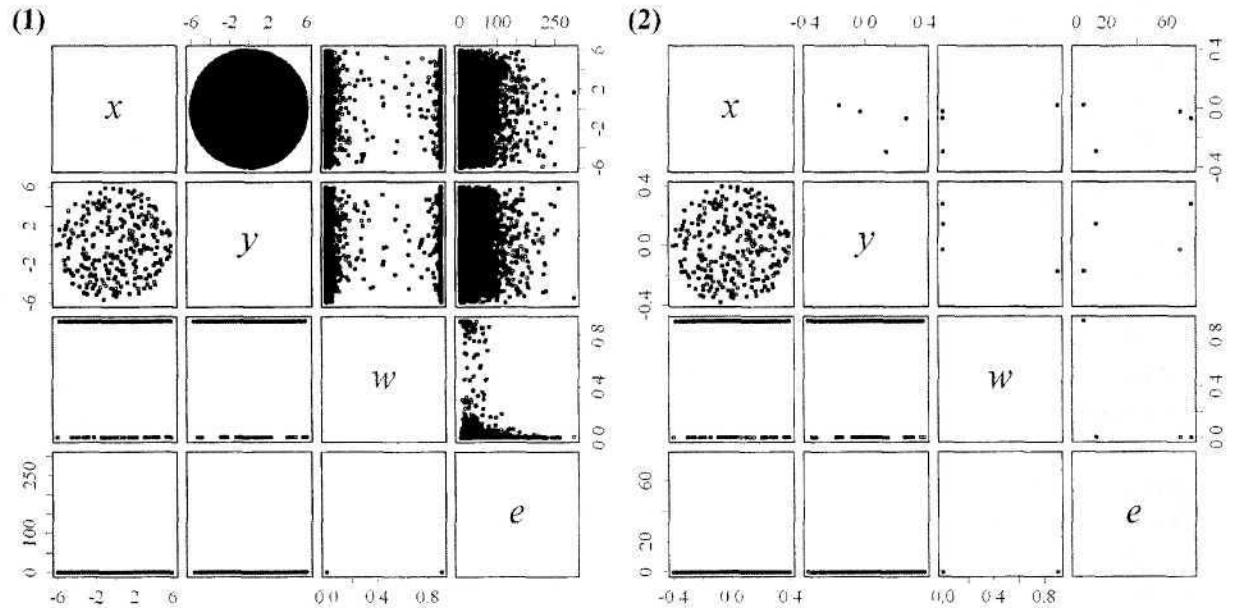


Fig. 10. The inclusion of a second diffractive element may improve significantly filtering and spatial distribution; (1) reference behavior and (2) predicted performance

In Fig. 10, we consider a 12 mm wide circular aperture, and diffractive elements in a 15X configuration; with  $G = 1200$  [lines/mm],  $m = +1$ , and  $\lambda = 632.8$  [nm]. Detection occurs immediately  $z = 30$  [mm] after the sample, case (1) in the figure, or after the diffractive elements, scenario (2). The optical properties of the 10 mm thick sample are those used throughout this work;  $\mu_a = 0$  [mm<sup>-1</sup>],  $\mu_s = 0.591$  [mm<sup>-1</sup>], and  $g = 0.93$ . The filtered distribution retains the circular pattern with a diameter of 800  $\mu\text{m}$ , but with-

out optical noise. Radiation from this coupler may be utilized as input to a BAI. By removing the asymmetry, the system may be employed for imaging and characterization of tissues with ballistic photons.

### 6. Conclusions and Future Work

In the ballistic auto-correlation interferometer (BAI), the arrangement is such that no reference contributions are in play. By developing the theory of the BAI, we predict that for-



ward-scattered and ballistic information coexists in the interferometric technique. Therefore, we emphasize the importance of implementing an efficient method to separate signal from undesired optical noise. This requirement is especially important for the assessment of biological samples. The latter assertion is a consequence of the forward-directed nature of scattering in tissues.

A non-linear grating-based angular filter may be employed to couple the information of interest to the BAI. The suggested system is conformed by a monochromatic source incident on a ruled grating, positioned at grazing diffraction, followed by a narrow slit. Depending on the angle of incidence of the transmitted radiation, the direction propagation of diffracted radiation is amplified non-linearly. In the experiment, amplification values ranging from 10-20X are feasible with 1 arc min resolution mounts. More sophisticated equipment may be employed to explore the highly non-linear regime of the coupler.

In transillumination experiments, such as the BAI, the features of angular amplification and decrease in beam extent are exploited to select the information of interest. Furthermore, the proposed system conforms an efficient ~100X coupling scheme. In contrast to traditional transillumination, the effects of forward-scattered contributions are minimized in the grating-based transmission experiment. In addition, greater exposition times (orders of magnitude) indicate detection of ballistic contributions in the grating-based schemes. Therefore, the proposed method may be suited for coupling selectively the radiation required in the BAI.

Our future research directions will be focused on the realization of a BAI. For this purpose, an important topic of investigation is the plausible implementation of a 2D grating-based coupler as the one discussed in this work. Such system should enable further improvements in filtering efficacy, and hence facilitate the coupling of radiation to the BAI.

*This work was carried out during an intership in the "International Institute of Optics and Biophotonics" at Saratov State University. P. Vacas-Jacques gratefully acknowledges the National Science and Technology Council of Mexico (CONACyT) for the financial support given*

*throughout the intership at Saratov State University, grant 172068. The same author thanks the "Centro de Investigaciones en Optica" for partial funding of travel expenses.*

*This research was performed in the framework of the project "Applications of infrared interferometry for biomedical tomography" (CONACyT 2007-10003-60450), and of the Russian Federation Program for the Development of High School Potential (2009-2010), projects № 2.1.1/4989 and № 2.2.1.1/2950.*

#### References

1. Patterson M., Wilson B., Wyman D. The propagation of optical radiation in tissue. I. Models of radiation transport and their application // *Lasers Med. Sci.* 1991. Vol.6. P.155-168.
2. Cubeddu R., Pifferi A., Taroni P., Torricelli A., Valentini G. Experimental test of theoretical models for time-resolved reflectance // *Med. Phys.* 1996. Vol.23. P.1625-1633.
3. Arridge S., Hebden J. Optical imaging in medicine: II. Modelling and reconstruction // *Phys. Med. Biol.* 1997. Vol.42. I.5. P.841-853.
4. Kim A. Transport theory for light propagation in biological tissue // *J. Opt. Soc. Amer. A.* 2004. Vol.21. I.5. P.820-827.
5. Patterson M., Wilson B., Wyman D. The propagation of optical radiation in tissue. II. Optical properties of tissues and resulting fluence distributions // *Lasers Med. Sci.* 1991. Vol.6. I.4. P.379-390.
6. Flock S., Wilson B., Patterson M. Total attenuation coefficients and scattering phase functions of tissues and phantom materials at 633 nm // *Med. Phys.* 1987. Vol.14. I.5. P.835-841.
7. Cheong W., Prahl S., Welch A. A review of the optical properties of biological tissues // *IEEE J. Quantum Elect.* 1990. Vol.26(12). P.2166-2185.
8. Лычагов В.В., Лякин Д.В., Модель М.Д., Рябухо В.П. Автореляционная низкокогерентная интерферометрия рассеивающих и слоистых объектов // *Компьютерная оптика.* 2007. Т.31, №3. С.40-51.
9. Modell M., Ryabukho V., Lyakin D., Lychagov V., Vitkin E., Itzkan I., Perelman L. Autocorrelation low coherence interferometry // *Opt. Commun.* 2008. Vol.281, №8. P.1991-1996.
10. Vacas-Jacques P., Paez G., Strojnik M. Pass-through photon-based biomedical transillumination // *J. Biomed. Opt.* 2008. Vol.13. I.4. P.301-307.
11. Vacas-Jacques P., Strojnik M., Paez G. Forward-calculated analytical interferograms in pass-through photon-based biomedical transillumination // *J. Opt. Soc. Amer. A* 2009. Vol.26. I.3. P.602-612.
12. Vacas-Jacques P., Ryabukho V., Strojnik M., Tuchin V., Paez G. Theoretical diffractive filter performance for ballistic transillumination // *Comp. Opt. / Компьютерная оптика.* 2009. Т.33, №2. С.129-137.
13. Sun K., Buchman S., Byer R. Grating angle magnification enhanced angular and integrated sensors for LISA applications // *J. Phys.: Conf. Ser.* 2006. Vol.32. I.1. P.167-179.



14. Sun K., Byer R. Grating angle magnification enhanced angular sensor and scanner: United States Patent №20080002197A1, Palo Alto (CA), 2008. Appl. №11/820. P.720.

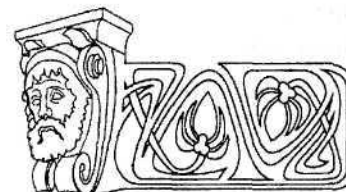
15. Vacas-Jacques P., Ryabukho V., Strojnik M., Tuchin V., Paez G. Non-linear grating-based angular filter for ballistic transillumination // Proc. SPIE. Novel Optical Instrumentation for Biomedical Applications / Ed. C. Depcursinge. Bellingham (WA), 2009. Vol.7371.

16. Pogue B., Patterson M. Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry // J. Biomed. Opt. 2006. Vol.11. I.4. P.096–102.

17. Fried D., Featherstone J., Darling C., Jones R., Ngaothepitak P., Bühler C. Early caries imaging and monitoring with near-infrared light // Dent. Clin. North Amer. 2005. Vol.49. I.4. P.771–793.

УДК 539.194

## ИК СПЕКТР МЕТИЛ-β-D-ГЛЮКОПИРАНОЗИДА И ЕГО ИНТЕРПРЕТАЦИЯ НА ОСНОВЕ ПОСТРОЕНИЯ СТРУКТУРНО-ДИНАМИЧЕСКОЙ МОДЕЛИ МОЛЕКУЛЫ



Л.М. Бабков, М.В. Королевич\*, Е.А. Моисейкина

Саратовский государственный университет  
E-mail: babkov@sgu.ru

\* Институт физики НАН Беларуси, Минск

Методом функционала плотности (B3LYP) в базисах 6-31 G(d), 6-31+G (d, p) построены структурно-динамические модели молекулы метил-β-D-глюкопиранозида: минимизированы энергии, рассчитаны структуры, дипольные моменты, поляризуемости, частоты нормальных колебаний в гармоническом приближении и распределение интенсивности в ИК спектре молекулы. Дана интерпретация ИК спектра метил-β-D-глюкопиранозида, измеренного в диапазоне 400–3700 см<sup>-1</sup> при комнатной температуре. Обсуждены преимущества построенной модели в сравнении с моделью, основанной на использовании метода валентно-силового поля и валентно-оптической теории.

**Ключевые слова:** метил-β-D-глюкопиранозид, молекулярное моделирование, метод функционала плотности, метод валентно-силового поля, валентно-оптическая теория, электрооптические параметры, механические параметры, нормальные колебания, ИК спектр, частота, интенсивность.

**IR Spectrum of the Methyl-β-D-Glucopyranoside and its Interpretation on Basis Structural-Dynamic Model Construction**

L.M. Babkov, M.V. Korolevich, E.A. Moisejkina

Structural-dynamic models of the methyl-β-D-glucopyranoside molecule are constructed by density functional method in bases 6-31 G(d), 6-31+G (d, p). Energies, structures, dipole moments, polarizabilities, frequencies of normal modes in harmonic approximation and IR intensities have been calculated. Interpretation of IR absorption spectrum is presented in range 400–3700 cm<sup>-1</sup>. Advantages of model, which was constructed, compared with model, which bases on using valence-force field method and valency-optical theory, are discussed.

**Key words:** methyl-β-D-glucopyranoside, molecular modeling, density functional method, valence-force field method, valence-optical theory, electro-optical parameters, mechanical parameters, normal modes, IR spectrum, frequencies, intensity.

### Введение

Моделирование структурно-динамических моделей молекул и молекулярных систем, составляющих вещество в различных

фазовых состояниях, на основе квантово-механических методов стало неотъемлемой частью исследований его структуры, оптических спектров и свойств. Точность квантово-механических расчетов за последние три десятилетия увеличилась на порядок, и это открыло возможность широкого их использования не только в квантовой химии, но и в молекулярной спектроскопии. Расхождение вычисленных частот нормальных колебаний и измеренных в спектрах комбинационного рассеяния (КР) или инфракрасного (ИК) поглощения в настоящее время не превышает 5%, что открыло возможность широкого использования результатов таких расчетов для полной интерпретации измеренных спектров. В пределах указанной погрешности абсолютная ошибка для частот нормальных колебаний оказывается различной. Для частот валентных колебаний связей С–Н и О–Н (высокочастотная область спектра, 2850–3750 см<sup>-1</sup>) она составляет 150–200 см<sup>-1</sup>. Для частот среднего диапазона (400–1600 см<sup>-1</sup>) ошибка на порядок меньше, что делает результаты расчетов в указанной области равноценными аналогичным результатам, получаемыми в рамках широко используемого до последнего времени в нашей стране метода валентно-силового поля теории колебательных спектров молекул [1–4]. В области ниже 400 см<sup>-1</sup>, в которой проявляются торсионные (кру-



тильные) и неплоские деформационные колебания, как показывает практика таких расчетов, расхождения снова возрастают.

Для надежной интерпретации измеренных спектров вещества, в особенности в случаях, когда исследуются соединения гомологического ряда, желательно снизить погрешность до минимума. Для этого обычно используют метод масштабирования силовых полей или частот [5–7]. Улучшения согласия опытных и рассчитанных частот можно ожидать при учете механического ангармонизма колебаний, как это показано для молекул бензола [8], азациклических соединений [9, 10].

Расчет распределения интенсивности в колебательных спектрах в рамках единого квантово-механического подхода дает значительно улучшенный результат по сравнению с подходом в рамках валентно-оптической теории [1–4]. Хорошее согласие рассчитанных интенсивностей колебаний структурных элементов молекулы, содержащих полярные связи (C=O, C–O, C≡N, O–H, C–X, где X – F, Cl, Br), с экспериментальными, которого не всегда удавалось добиться в рамках валентно-оптической теории, позволяет более обоснованно интерпретировать спектры.

Объектом исследования, результаты которого излагаются в данной статье, стал метил-β-D-глюкопиранозид. Метил-β-D-глюкопиранозид принадлежит классу соединений, близких к биообъектам. Это углеводы (моносахариды), свойства которых определены их строением, внутри- и межмолекулярными взаимодействиями. По колебательным спектрам таких соединений в Институте физики НАН Беларуси группой Р.Г. Жбанкова накоплен и систематизирован обширный экспериментальный материал. Колебательные спектры многих из них интерпретированы на основе результатов расчетов по оригинальным программам, в которых реализованы методы валентно-силового поля и валентно-оптической теории [1–4].

В работе [12] ИК спектр метил-β-D-глюкопиранозид, измеренный при комнатной температуре, теоретически интерпретирован на основе результатов расчета упомянутыми методами для изолированной молекулы в области 880–1500 см<sup>-1</sup>. В квантово-химической части исследований использован ме-

тод CNDO/2. Влияние водородной связи в модели изолированной молекулы учтено эффективно.

В настоящей статье приведены и обсуждаются результаты проведенного авторами теоретического исследования структуры и колебательного спектра молекулы метил-β-D-глюкопиранозид в гармоническом приближении на основе современного квантово-механического метода теории функционала плотности (B3LYP), реализованного в комплексе программ GAUSSIAN'03 для Windows [11]. Целью исследований стала проверка используемых методов и приближений применительно к метил-β-D-глюкопиранозиду и предварительная интерпретация на основе результатов моделирования его ИК спектра. Полученные результаты сравниваются с результатами работы [12].

## 1. Моделирование

Моделирование проведено методом теории функционала плотности (B3LYP) в базисе 6-31+G (d). Базис включает одну функцию для внутренней оболочки, скомбинированную из шести гауссовых функций, и набор из внутренней функции, скомбинированной из трех гауссовых, и внешней не-скомбинированной гауссовой функции для валентной оболочки каждого атома. Кроме этого для увеличения точности расчета в базис было включено по одному набору не-скомбинированных d-функций на каждый тяжелый атом [13]. Был использован и более широкий базис 6-31+G (d, p), с добавлением по одному набору не-скомбинированных p-функций на каждый атом водорода в надежде на получение улучшенного результата.

Построены два варианта структурно-динамической модели молекулы: рассчитаны энергия, структура, дипольный момент, поляризуемость, частоты нормальных колебаний в гармоническом приближении и распределение интенсивности в ИК спектре свободной молекулы.

## 2. Результаты и их обсуждение

В результате минимизации энергии молекулы в базисе 6-31G (d) получено значение энергии, равное -726.45227 Hartree. Результаты оптимизации геометрии приведены на

рис. 1. Для удобства при анализе результатов расчетов геометрии молекулы и интерпретации ее спектра введены индексы: R – атомам шестичленного кольца; R1 – атомам гидроксильной группы OH; R2 – атомам группы  $\text{CH}_2\text{-OH}$ ; R3 – атомам группы  $\text{O-CH}_3$ .

Рассчитанные геометрические параметры молекулы приведены в табл. 1. Молекула не обладает симметрией. Неплоское кольцо имеет конформацию «кресло». Общий дипольный момент и его проекции на оси координат составили:  $\mu_0 = 2.8067$ ,  $\mu_x = -2.8059$ ;  $\mu_y = -0.0588$ ;  $\mu_z = 0.0250$  Дебая.

На рис. 2 приведен рассчитанный спектр молекулы метил-β-D-глюкопиранозида (2), где он сравнивается с экспериментальным спектром (1), измеренным в работе [12]. Интерпретация ИК спектра метил-β-D-глюкопиранозида дана в табл. 2 на основе отношения частот нормальных колебаний по их формам и интенсивностям.

Рассчитанные в базисе 6-31+G (d, p) строение и геометрия молекулы не отличаются от полученных в базисе 6-31G (d), за исключением небольших различий в углах между связями в кольце. Электрооптические параметры также практически совпадают. Вычисленные частоты несколько хуже согласуются с измеренными: они ещё более завышены в высокочастотной области. В низкочастотной области (меньше  $400 \text{ см}^{-1}$ ) они занижены по сравнению с полученными в базисе 6-31G (d). Таким образом, улучшение базиса за счет добавления в него *p*-функций не увеличило точность расчета для данного соединения, поэтому мы опирались на результаты, полученные с использованием базиса 6-31G (d), не анализируя на данном этапе причин сложившейся ситуации, приняв ее как факт.

В высокочастотной области ИК спектра ( $2850\text{--}3700 \text{ см}^{-1}$ ) вычисленные значения частот валентных колебаний связей C–H и O–H оказались завышенными по сравнению с экспериментальными на 4–5%. Отчасти эти расхождения обусловлены ограниченностью гармонического приближения. В работе [12] значения этих частот ближе к экспериментальным. В нашем случае расхождение было устранено масштабированием этих частот.

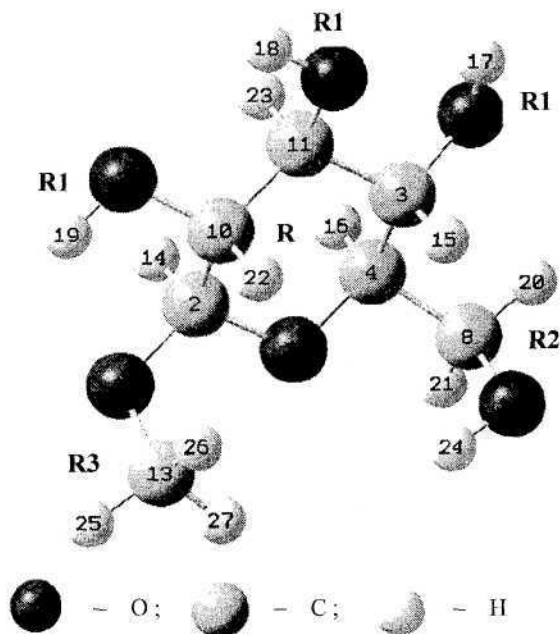


Рис. 1. Строение молекулы метил-β-D-глюкопиранозида

Таблица 1  
Рассчитанные геометрические параметры молекулы метил-β-D-глюкопиранозида

Длины связей, Å					
Связь	Длина связи	Связь	Длина связи	Связь	Длина связи
O <sub>1</sub> -C <sub>2</sub>	1.417	C <sub>2</sub> -O <sub>9</sub>	1.402	O <sub>12</sub> -H <sub>24</sub>	0.971
O <sub>1</sub> -C <sub>4</sub>	1.439	O <sub>9</sub> -C <sub>13</sub>	1.428	C <sub>2</sub> -H <sub>14</sub>	1.101
C <sub>3</sub> -C <sub>4</sub>	1.535	C <sub>13</sub> -H <sub>25</sub>	1.092	C <sub>3</sub> -H <sub>15</sub>	1.100
C <sub>2</sub> -C <sub>10</sub>	1.536	C <sub>13</sub> -H <sub>26</sub>	1.098	C <sub>4</sub> -H <sub>16</sub>	1.102
C <sub>3</sub> -C <sub>11</sub>	1.526	C <sub>13</sub> -H <sub>27</sub>	1.094	C <sub>10</sub> -H <sub>22</sub>	1.101
C <sub>10</sub> -C <sub>11</sub>	1.522	C <sub>4</sub> -C <sub>8</sub>	1.527	C <sub>11</sub> -H <sub>23</sub>	1.103
C <sub>10</sub> -O <sub>7</sub>	1.422	C <sub>8</sub> -O <sub>12</sub>	1.418	O <sub>5</sub> -H <sub>17</sub>	0.972
C <sub>11</sub> -O <sub>6</sub>	1.423	C <sub>8</sub> -H <sub>20</sub>	1.093	O <sub>6</sub> -H <sub>18</sub>	0.971
C <sub>3</sub> -O <sub>5</sub>	1.419	C <sub>8</sub> -H <sub>21</sub>	1.101	O <sub>7</sub> -H <sub>19</sub>	0.972
Углы между связями, °					
Угол	Величина угла	Угол	Величина угла	Угол	Величина угла
C <sub>2</sub> -O <sub>1</sub> -C <sub>4</sub>	114.2	O <sub>1</sub> -C <sub>4</sub> -H <sub>16</sub>	109.6	H <sub>25</sub> -C <sub>13</sub> -H <sub>27</sub>	109.5
O <sub>1</sub> -C <sub>2</sub> -C <sub>10</sub>	109.8	C <sub>2</sub> -C <sub>10</sub> -O <sub>7</sub>	111.2	H <sub>26</sub> -C <sub>13</sub> -H <sub>27</sub>	109.4
O <sub>1</sub> -C <sub>2</sub> -O <sub>9</sub>	110.1	C <sub>11</sub> -C <sub>10</sub> -O <sub>7</sub>	107.4	C <sub>4</sub> -C <sub>8</sub> -H <sub>20</sub>	109
O <sub>9</sub> -C <sub>2</sub> -H <sub>14</sub>	104.6	C <sub>2</sub> -C <sub>10</sub> -H <sub>22</sub>	109.7	C <sub>4</sub> -C <sub>8</sub> -H <sub>21</sub>	108.7
C <sub>10</sub> -C <sub>2</sub> -H <sub>14</sub>	109.2	C <sub>11</sub> -C <sub>10</sub> -H <sub>22</sub>	108.8	O <sub>12</sub> -C <sub>8</sub> -H <sub>20</sub>	106.8
C <sub>4</sub> -C <sub>3</sub> -C <sub>11</sub>	110.2	C <sub>3</sub> -C <sub>11</sub> -O <sub>6</sub>	107.3	O <sub>12</sub> -C <sub>8</sub> -H <sub>21</sub>	110.9
C <sub>4</sub> -C <sub>3</sub> -H <sub>15</sub>	108.9	C <sub>3</sub> -C <sub>11</sub> -C <sub>10</sub>	110.7	C <sub>8</sub> -O <sub>12</sub> -H <sub>24</sub>	107.5
O <sub>5</sub> -C <sub>3</sub> -C <sub>11</sub>	110.9	C <sub>10</sub> -C <sub>11</sub> -H <sub>23</sub>	108.4	C <sub>3</sub> -O <sub>5</sub> -H <sub>17</sub>	107.4
O <sub>5</sub> -C <sub>3</sub> -H <sub>15</sub>	110.4	O <sub>6</sub> -C <sub>11</sub> -H <sub>23</sub>	109.6	C <sub>11</sub> -O <sub>6</sub> -H <sub>18</sub>	108.1
C <sub>3</sub> -C <sub>4</sub> -C <sub>8</sub>	113.5	C <sub>2</sub> -O <sub>9</sub> -C <sub>13</sub>	117.2	C <sub>10</sub> -O <sub>7</sub> -H <sub>19</sub>	108.4
O <sub>1</sub> -C <sub>4</sub> -C <sub>8</sub>	106.3	O <sub>9</sub> -C <sub>13</sub> -H <sub>25</sub>	105.5	–	–
C <sub>3</sub> -C <sub>4</sub> -H <sub>16</sub>	108.7	O <sub>6</sub> -C <sub>13</sub> -H <sub>26</sub>	112.3	–	–

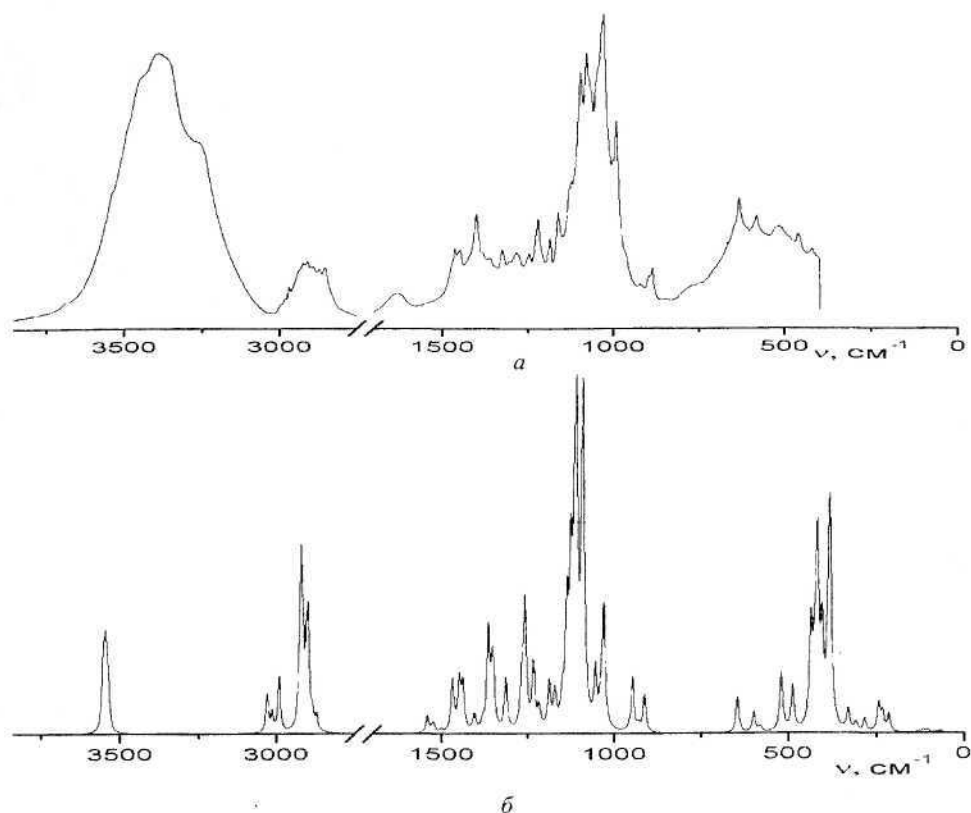


Рис. 2. ИК спектр метил-β-D-глюкопиранозида: а – эксперимент, б – теория

Таблица 2

Измеренные и рассчитанные частоты и интенсивности фундаментальных колебаний ИК спектра метил-β-D-глюкопиранозида

№	$\nu_{\text{эксп}}(\text{I}), \text{cm}^{-1}$	$\nu_{\text{теор}}, \text{cm}^{-1}$	I, км/моль	Форма колебания
1	–	67	1,7	$\chi_{\text{R}}(\text{HCCH}), \chi_{\text{R}}(\text{CCCH}), \chi_{\text{RR2}}(\text{HCCC}), \chi_{\text{RR2}}(\text{CCCC})$
2	–	78	0,7	$\chi_{\text{RR3}}(\text{OCOC}), \chi_{\text{RR3}}(\text{CCOC}), \chi_{\text{RR3}}(\text{HCOC}), \chi_{\text{RR3}}(\text{COCH})$
3	–	103	2,5	$\chi_{\text{RR3}}(\text{OCOC}), \chi_{\text{RR3}}(\text{CCOC}), \chi_{\text{RR3}}(\text{HCOC}), \chi_{\text{RR3}}(\text{COCH})$
4	–	116	2,5	$\chi_{\text{RR1}}(\text{HCCO}), \chi_{\text{R}}(\text{HCCH}), \chi_{\text{R1R}}(\text{OCCH}), \chi_{\text{R}}(\text{CCCH})$
5	–	132	0,9	$\chi_{\text{RR2}}(\text{CCCO}), \chi_{\text{RR2}}(\text{CCCH}), \chi_{\text{RR2}}(\text{HCCO}), \chi_{\text{RR2}}(\text{HCCH})$
6	–	170	0,4	$\chi_{\text{RR3}}(\text{COCH})$
7	–	214	12,2	$\chi_{\text{R1R}}(\text{HOCC}), \chi_{\text{R1R}}(\text{HOCH}), \chi_{\text{R1R}}(\text{OCCO}), \chi_{\text{RR1}}(\text{CCOH})$
8	–	232	13,0	$\chi_{\text{R1R}}(\text{HOCC}), \chi_{\text{R1R}}(\text{HOCH}), \chi_{\text{RR3}}(\text{COCH}), \chi_{\text{RR2}}(\text{OCCH})$
9	–	243	19,2	$\chi_{\text{RR1}}(\text{CCOH}), \chi_{\text{RR3}}(\text{COCH}), \chi_{\text{RR1}}(\text{HCOH}), \chi_{\text{R1R}}(\text{HOCC})$
10	–	257	0,5	$\chi_{\text{RR2}}(\text{CCCH}), \chi_{\text{RR2}}(\text{COCC}), \chi_{\text{RR2}}(\text{CCCC}), \chi_{\text{R}}(\text{COCC})$
11	–	283	8,9	$\chi_{\text{RR3}}(\text{COCH}), \chi_{\text{RR2}}(\text{OCCH}), \chi_{\text{RR3}}(\text{COCO}), \chi_{\text{RR3}}(\text{CCOC})$
12	–	309	6,2	$\chi_{\text{R1R}}(\text{HOCC}), \chi_{\text{R1R}}(\text{HOCH}), \chi_{\text{RR3}}(\text{OCOC}), \chi_{\text{RR3}}(\text{HCOC})$
13	–	330	14,3	$\chi_{\text{RR1}}(\text{CCOH}), \chi_{\text{RR1}}(\text{HCOH}), \chi_{\text{RR3}}(\text{COCH}), \chi_{\text{RR3}}(\text{OCOC})$
14	–	382	135,4	$\chi_{\text{R1R}}(\text{HOCC}), \chi_{\text{R1R}}(\text{HOCH}), \chi_{\text{RR1}}(\text{CCOH}), \chi_{\text{RR1}}(\text{HCOH})$
15	–	387	50,2	$\chi_{\text{R1R}}(\text{HOCC}), \chi_{\text{R1R}}(\text{HOCH}), \chi_{\text{RR1}}(\text{CCOH}), \chi_{\text{RR1}}(\text{HCOH})$
16	–	405	60,6	$\chi_{\text{R1R}}(\text{HOCC}), \chi_{\text{R1R}}(\text{HOCH}), \chi_{\text{RR1}}(\text{CCOH}), \chi_{\text{RR1}}(\text{HCOH})$
17	–	418	122,9	$\chi_{\text{RR1}}(\text{CCOH}), \chi_{\text{RR1}}(\text{HCOH}), \chi_{\text{RR2}}(\text{CCOH}), \chi_{\text{R2}}(\text{HCOH})$
18	–	426	33,7	$\chi_{\text{R1R}}(\text{HOCH}), \chi_{\text{R1R}}(\text{HOCC}), \chi_{\text{R2}}(\text{HCOH}), \chi_{\text{RR2}}(\text{CCOH})$
19	–	436	67,8	$\chi_{\text{P P2}}(\text{XXOH}), \chi_{\text{P2}}(\text{HXOH}), \chi_{\text{P P1}}(\text{XXOH}), \chi_{\text{P P1}}(\text{HXOH})$
20	–	488	30,6	$\chi_{\text{P2}}(\text{HXOH}), \chi_{\text{P P2}}(\text{XXOH}), \chi_{\text{P P1}}(\text{XXOH}), \chi_{\text{P P1}}(\text{HXOH})$
21	518(ср)	520	40,1	$\chi_{\text{P1 P}}(\text{HOXX}), \chi_{\text{P1 P}}(\text{OXXO}), \chi_{\text{P1 P}}(\text{HOXH}), \chi_{\text{P1 P}}(\text{HXXO})$



№	$\nu_{\text{эксп}}(\text{I}), \text{см}^{-1}$	$\nu_{\text{теор}}, \text{см}^{-1}$	I, км/моль	Форма колебания
22	583(ср)	580	3,9	$\chi_{R2}$ (HCOH), $\chi_{RR2}$ (CCOH), $\chi_{RR2}$ (OCCH), $\gamma_{R1R}$ (OCC)
23		598	13,8	$\chi_R$ (COCC), $\chi_{RR1}$ (HCCO), $\chi_{RR3}$ (OCOC), $\gamma_R$ (OCC)
24	635(ср)	645	25,3	$\chi_{R2}$ (HCOH), $\chi_{RR2}$ (CCOH), $\gamma_{RR3}$ (COC), $\chi_{R1R}$ (OCCO)
25	—	728	0,4	$\gamma_{RR2}$ (CCO), $\chi_R$ (CCCC), $\gamma_{RR2}$ (OCC), $\chi_{RR1}$ (OCCO)
26	890(сл)	913	25,7	$\chi_R$ (CCCH), $\beta_{R2R}$ (CCH), $\chi_R$ (COCH), $\chi_R$ (CCCC)
27	—	930	3,5	$\chi_{R2}$ (HCOH), $\beta_{RR2}$ (CCH), $\chi_{RR2}$ (OCCH), $\chi_{RR2}$ (CCCH)
28	—	947	37,6	$Q_{R3}$ (OC), $Q_{R1R}$ (OC), $\chi_{R1R}$ (HOCH), $\chi_R$ (COCC)
29	1030(о.с)	1030	80,7	$Q_{R1R}$ (OC), $\beta_{RR1}$ (COH), $Q_{R2}$ (CO), $\chi_R$ (HCCH), $Q_{R3}$ (OC)
30		1038	19,2	$\beta_{RR1}$ (COH), $Q_{R3}$ (OC), $\chi_{RR2}$ (HCCC), $\chi_R$ (HCCH)
31	1053(с)	1055	38,4	$Q_{R2}$ (CO), $Q_R$ (CC), $\chi_{RR2}$ (OCCH), $\chi_{R1R2}$ (OCCC)
32	1081(с)	1079	5,6	$\chi_R$ (CCCC), $\chi_R$ (OCCC), $\chi_R$ (COCH), $\chi_{R3R1}$ (OCCO)
33		1092	226,1	$Q_R$ (CC), $\beta_{R2}$ (COH)
34	1098(о.с)	1110	180,0	$\beta_{R2}$ (COH), $\beta_{R2}$ (OCH), $\chi_{RR2}$ (CCCH), $Q_{R1R}$ (OC)
35		1116	80,2	$Q_{R2}$ (CO), $Q_R$ (CC), $\chi_R$ (CCCC), $\chi_R$ (CCCH), $\chi_R$ (CCCC)
36	1124(ср)	1126	100,9	$Q_R$ (CC), $\beta_{RR1}$ (COH), $\gamma_R$ (CCC)
37		1137	76,5	$Q_{R1R}$ (OC), $\beta_{RR1}$ (COH), $Q_{R2}$ (CO), $\gamma_{R3R}$ (OCC), $Q_{R3}$ (OC)
38		1151	19,2	$Q_{RR1}$ (CO), $Q_{R1R}$ (OC), $Q_R$ (CC), $\gamma_R$ (CCC)
39	1163(ср)	1172	23,8	$\beta_{R3}$ (OCH), $Q_{R3}$ (OC), $\chi_{RR3}$ (COCH), $\chi_R$ (HCCC)
40	1187(ср)	1189	31,0	$\beta_{R3}$ (OCH), $\chi_{RR3}$ (COCH), $\chi_{R1R}$ (HOCH), $\beta_R$ (CCH)
41	1222(ср)	1218	14,1	$\chi_{RR2}$ (HCCH), $\beta_R$ (CCH), $\chi_{R1R}$ (HOCH), $\chi_{RR1}$ (HCOH)
42		1235	44,3	$\beta_{R3}$ (OCH), $\chi_{RR3}$ (COCH), $Q_{RR3}$ (CO)
43	1240(ср)	1259	22,9	$\chi_R$ (HCCH), $\beta_R$ (CCH), $\beta_{RR2}$ (CCH), $\chi_{R1R}$ (HOCH)
44		1259	63,9	$\chi_R$ (OCCH), $\beta_R$ (CCH), $\chi_R$ (HCCH)
45	1249(ср)	1269	30,5	$\chi_R$ (HCCH), $\beta_R$ (CCH), $\chi_R$ (HCCC), $\chi_{RR1}$ (HCOH)
46	1305(ср)	1314	35,7	$\chi_R$ (HCCH), $\chi_{RR3}$ (HCOC), $\chi_R$ (HCCC), $\beta_R$ (OCH)
47	1325(сл)	1350	22,7	$\chi_{R1R}$ (OCCH), $\chi_R$ (HCCH), $\beta_{R2R}$ (CCH), $\chi_R$ (CCCH)
48	1352(сл)	1353	29,7	$\chi_{RR2}$ (HCCH), $\beta_R$ (CCH), $\chi_{RR2}$ (HCCO), $\beta_R$ (OCH)
49	1362(ср)	1364	69,7	$\chi_R$ (HCCH), $\beta_{RR1}$ (COH), $\chi_R$ (CCCH), $\chi_{R1R}$ (OCCH)
50		1381	5,6	$\chi_R$ (HCCH), $\chi_R$ (HCCO), $\chi_{RR2}$ (HCCC)
51	1401(ср)	1403	10,5	$\beta_{R2}$ (OCH), $\beta_{RR2}$ (CCH), $\beta_R$ (OCH)
52	—	1410	1,6	$\chi_R$ (CCCH), $\chi_{R1R}$ (OCCH), $\chi_{RR2}$ (HCCC), $\chi_R$ (HCCO)
53	1448(сл)	1437	13,6	$\chi_R$ (HCCH), $\chi_R$ (COCH), $\chi_{RR1}$ (HCCO), $\beta_{R3R}$ (OCH)
54		1438	17,3	$\beta_{R3R}$ (OCH), $\chi_{RR1}$ (HCCO), $\chi_R$ (COCH), $\chi_R$ (HCCC)
55		1448	33,6	$\beta_{R2}$ (COH), $\beta_{R2}$ (OCH), $\beta_{RR2}$ (CCH), $\chi_{RR2}$ (OCCH)
56	1464(сл)	1465	4,1	$\chi_{R1R}$ (OCCH), $\beta_{R1R}$ (OCH), $\chi_R$ (HCCH), $\chi_R$ (CCCH)
57		1467	31,6	$\chi_{RR1}$ (HCCO), $\chi_R$ (HCCC), $\beta_R$ (CCH), $\beta_{RR1}$ (COH)
58	—	1499	2,1	$\alpha_{R3}$ (HCH), $\beta_{R3}$ (OCH)
59	—	1520	3,5	$\alpha_{R3}$ (HCH), $\chi_{RR3}$ (COCH)
60	—	1525	4,6	$\alpha_{R2}$ (HCH), $\chi_{RR2}$ (CCCH), $\chi_{RR2}$ (HCCH), $\chi_{RR2}$ (OCCH)
61	—	1540	11,3	$\alpha_{R3}$ (HCH), $\chi_{RR3}$ (COCH)
62	—	2871	9,3	$q_R$ (CH)
63	—	2885	10,2	$q_{R2}$ (CH)
64	2893(ср)	2890	0,1	$q_R$ (CH)
65		2899	65,0	$q_{R2}$ (CH)
66		2904	34,7	$q_R$ (CH)



№	$\nu_{\text{эксп}}(\text{I}), \text{см}^{-1}$	$\nu_{\text{теор}}, \text{см}^{-1}$	$I, \text{км/моль}$	Форма колебания
67	2911(ср)	2916	56,0	$q_{R3}$ (СН)
68		2922	103,7	$q_{R3}$ (СН)
69	–	2991	38,5	$q_{R3}$ (СН)
70	–	3013	13,0	$q_{R2}$ (СН)
71	–	3030	26,0	$q_{R3}$ (СН)
72	3375(о.с.)	3537	31,5	$q_{R1}$ (ОН)
73		3544	39,6	$q_{R1}$ (ОН)
74		3547	17,4	$q_{R2}$ (ОН)
75		3552	40,7	$q_{R1}$ (ОН)

В области  $800\text{--}1500 \text{ см}^{-1}$ , где проявляются валентные и деформационные колебания шестичленного кольца и связей С–О, достигнуто лучшее согласие рассчитанных и измеренных частот: среднее абсолютное отклонение между ними составляет  $8 \text{ см}^{-1}$  (в работе [12] –  $7 \text{ см}^{-1}$ ), и полученные в данной области результаты практически не нуждаются в корректировке.

Из-за отсутствия экспериментальных данных по ИК спектру в области ниже  $400 \text{ см}^{-1}$  предметно обсудить результаты расчета частот в этой области не представляется возможным.

Из рис. 2 видно, что рассчитанные частоты фундаментальных колебаний и их интенсивности находятся в удовлетворительном согласии с экспериментальными. Результаты по интенсивностям превосходят полученные в работе [12]. Исключением стали частоты и интенсивности валентных колебаний связей О–Н (область  $3200\text{--}3750 \text{ см}^{-1}$ ).

Анализ области  $3200\text{--}3750 \text{ см}^{-1}$  измеренного спектра указывает на сильное расхождение в значениях экспериментальных и теоретических частот и интенсивностей валентных колебаний связей О–Н. В рассчитанном спектре эти колебания оказываются в узком интервале  $3552\text{--}3537 \text{ см}^{-1}$  (см. табл. 2, рис. 2). Центр тяжести полосы измеренного спектра соответствует  $3310 \text{ см}^{-1}$ . Таким образом, можно констатировать наличие сдвига частотных положений полос экспериментального спектра, соответствующих этим колебаниям, порядка  $240 \text{ см}^{-1}$ . Этот сдвиг является одним из главных признаков образования в образце водородных связей. Аномаль-

но большие интенсивность, ширина, сложная форма полосы, соответствующей указанным колебаниям, в измеренном спектре также указывают на образование в образце комплексов с водородной связью. Теоретический же спектр отражает динамику свободной молекулы, поэтому сильно отличается в данной области от экспериментального.

Вопрос о влиянии водородной связи на структуру и спектр метил- $\beta$ -D-глюкопиранозиды представляется чрезвычайно важным с точки зрения теории строения и межмолекулярных взаимодействий в приложении к данному соединению и заслуживает отдельного рассмотрения.

#### Выводы

В результате построения структурно-динамической модели свободной молекулы метил- $\beta$ -D-глюкопиранозиды определены ее геометрические параметры. Молекула не обладает симметрией, ее неплоское кольцо имеет конформацию «кресло».

Сдвиг рассчитанного ИК спектра в области валентных колебаний связей С–Н и О–Н относительно экспериментального в высокочастотную сторону, составляющий 4–5%, обусловлен ограниченностью гармонического приближения и может быть устранен масштабированием частот.

Большие расхождения по частотам и интенсивностям валентных колебаний связей О–Н в рассчитанном и измеренном ИК спектрах обусловлены влиянием водородной связи, которое при построении структурно-динамической модели свободной молекулы не учитывалось.





Добавление  $p$ -функций в базис 6–31G(d) не привело к улучшению точности расчетов для молекулы метил- $\beta$ -D-глюкопиранозида.

Сравнение результатов проведенного моделирования для молекулы метил- $\beta$ -D-глюкопиранозида с результатами работы [12] показало преимущество использованного подхода, основанного на методе теории функционала плотности, реализованного в программном комплексе GAUSSIAN<sup>03</sup>, перед традиционным подходом, использующим метод валентно-силового поля и валентно-оптической теории. На данном этапе исследования это преимущество проявилось в оценке распределения интенсивности в ИК спектре. Есть надежда, что в дальнейших исследованиях структуры метил- $\beta$ -D-глюкопиранозида и его ИК спектра потенциальные возможности использованного метода проявятся в большей степени.

#### Список литературы

1. Волькенштейн М.А., Грибов Л.А., Еляшевич М.А., Степанов Б.И. Колебания молекул. М.: Наука, 1972. 700 с.
2. Свердлов Л.М., Ковнер М.А., Крайнов Е.П. Колебательные спектры многоатомных молекул. М.: Наука, 1970. 560 с.
3. Грибов Л. А. Введение в молекулярную спектроскопию. М.: Наука, 1976. 400 с.
4. Грибов Л. А., Дементьев В. А. Моделирование колебательных спектров сложных соединений на ЭВМ. М.: Наука, 1989. 160 с.
5. Pulay P., Fogarasi G., Pongor G., Boggs J.E., Vargha A. Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. Scaled quantum mechanical (QM) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene // J. Amer. Chem. Soc. 1983. Vol.105. P.7037–7047.
6. Yashida H., Takeda K., Okamura J., Ehara A., Matsui H. A New Approach to Vibrational Analysis of Large Molecules by Density Functional Theory: Wavenumber-Linear Scaling Method // J. Phys. Chem. A. 2002. Vol.106, №14. P.3580–3586.
7. Березин К.В., Кривохижина Т.В., Нечаев В.В. Применение метода линейного масштабирования частот в расчетах нормальных колебаний многоатомных молекул // Оптика и спектроскопия. 2003. Т.94, №3. С.398–401.
8. Willets A., Handy N.C. Anharmonic constants for benzene // Spectrochim. Acta. 1997. Vol.53, №8. P.1169–1177.
9. Березин К.В., Нечаев В.В., Элькин П.М. Ангармонический анализ колебательных состояний пиримидина методом функционала плотности // Оптика и спектроскопия. 2004. Т.97, №2. С.224–234.
10. Березин К.В., Нечаев В.В., Элькин П.М. Ангармонические резонансы в колебательных спектрах пиридазина // Журн. физ. химии. 2005. Т.79, №3. С.1–10.
11. Frisch J., Trucks G.W., Schlegel H.B. Gaussian03, Revision B.03; Gaussian, Inc., Pittsburgh PA, 2003. 302 p.
12. Королевич М.В., Жбанкова Р.В. Теоретическая интерпретация спектральных признаков оксиметильной группы в ИК спектре метил- $\beta$ -D-глюкопиранозида // Журн. прикл. спектроскопии. 2006. Т.73, №6. С.721–727.
13. Попл Дж.А. Квантово-химические модели // Успехи физ. наук. 2002. Т.172. С.349–356.

УДК 539.194

## КОЛЕБАТЕЛЬНЫЕ СПЕКТРЫ 2-БИФЕНИЛМЕТАНОЛА И ИХ ИНТЕРПРЕТАЦИЯ НА ОСНОВЕ МОДЕЛИРОВАНИЯ МОЛЕКУЛЫ МЕТОДОМ ТЕОРИИ ФУНКЦИОНАЛА ПЛОТНОСТИ С УЧЕТОМ МЕХАНИЧЕСКОГО АНГАРМОНИЗМА



Л.М. Бабков, К.Е. Успенский, Н.А. Давыдова\*

Саратовский государственный университет

E-mail: babkov@sgu.ru

\* Институт физики НАН Украины, Киев

E-mail: davydova@iop.kiev.ua

В приближении гибридного функционала плотности B3LYP/6–31G(d) произведен расчет геометрии, электрооптических параметров, квартичного силового поля и колебательных спектров молекулы 2-бифенилметанола. Проведен колебательный анализ и дана полная интерпретация его спектров ИК поглощения ( $400\text{--}3800\text{ см}^{-1}$ ) и комбинационного рассеяния ( $50\text{--}3600\text{ см}^{-1}$ ).

**Ключевые слова:** 2-бифенилметанол, ИК спектр, спектр КР, метод функционала плотности, молекулярное моделирование, механический ангармонизм, квартичное силовое поле, фундаментальная частота, составная частота, обертоны.



### The Vibrational Spectra of 2-Biphenylmethanol and their Interpretation by Using of Molecular Structural-Dynamical Model Taking into Account Mechanical Anharmonicity

L.M. Babkov, K.E. Uspenskiy, N.A. Davydova

Using hybrid density functional (B3LYP/6-31(d)) the calculation of the geometry, electro-optical parameters, quartic force field and vibrational spectra of the 2-biphenylmethanol molecule were performed. The vibrational analysis was implemented and the full interpretation of its IR (400–3800  $\text{cm}^{-1}$ ) and Raman (50–3600  $\text{cm}^{-1}$ ) spectra was given.

**Key words:** 2-biphenylmethanol, IR spectrum, Raman spectrum, density functional method, molecular modeling, mechanical anharmonicity, quartic force field, fundamental frequency, combinational frequency, overtone.

#### Введение

Хорошее согласие вычисленных и измеренных спектроскопических параметров является критерием правильности выбранных квантово-механических методов и приближений. Этому критерию, в частности, удовлетворяет метод теории функционала плотности (ТФП) B3LYP, широко используемый в мировой практике при построении структурно-динамических моделей молекулярных систем и реализованный в современных комплексах программ GAUSSIAN [1]. Расхождение вычисленных частот нормальных колебаний с измеренными не превышает 4–5%, но по абсолютному значению оно не одинаково для различных областей спектра: для валентных колебаний связей С–Н оно достигает  $150 \text{ см}^{-1}$ , для других колебаний в диапазоне  $400\text{--}1600 \text{ см}^{-1}$  оно находится в пределах, сопоставимых с результатами, получаемыми методом валентно-силового поля [2–5]. Для минимизации расхождения в мировой практике широко используют метод масштабирования силовых полей или частот [6–8]. Минимизировать расхождения можно при учете механического ангармонизма колебаний. Но этот учет требует значительного увеличения временного ресурса. Нельзя предугадать, приведет ли он к желаемому результату в низкочастотном диапазоне колебательного спектра (ниже  $400 \text{ см}^{-1}$ ). До настоящего времени учет ангармонизма проводился лишь для свободных молекул не слишком громоздких и хорошо изученных соединений: бензол [9], азациклические соединения [10, 11].

При расчете интенсивностей нормальных колебаний в спектрах на основе метода ТФП получается улучшенный результат по сравнению с подходом в рамках валентно-оптической теории, в основном по причине дефицита необходимых электрооптических параметров, в которых она нуждается. В первую очередь это относится к колебаниям структурных элементов молекул, содержащих полярные связи. При использовании удовлетворительных данных по интенсивностям интерпретации измеренных спектров становится более обоснованной.

В настоящей работе обсуждены результаты моделирования структуры молекулы 2-бифенилметанола (2БФМ) в рамках метода B3LYP в базисе 6-31G(d) и расчета его колебательных спектров с учетом механического ангармонизма, который, по определению, должен улучшить согласие рассчитанных и измеренных частот колебательных спектров без использования физически менее обоснованной процедуры масштабирования частот и обеспечить полную интерпретацию экспериментальных спектров, в которых наряду с нормальными колебаниями имеются их обертоны, составные колебания и, возможно, резонансы. Это составляет преимущество подхода, основанного на учете ангармонизма колебаний.

2БФМ – новое, недавно синтезированное вещество, удобное как объект для изучения стеклофазы, которой оно обладает. Механизм стеклования ещё не до конца понят [12]. При экспериментальных исследованиях статических и динамических свойств в перерожденном жидком состоянии основным требованием к веществу является его высокая термическая устойчивость по отношению к кристаллизации. Обладая высокой способностью к стеклованию, 2БФМ удовлетворяет этому требованию. Исследование микроструктуры стекол – одно из важных фундаментальных направлений в физике конденсированного состояния, так как она определяет особенности перехода вещества в аморфное состояние [13] и динамику составляющих его квантовых молекулярных систем. Комплексное исследование спектров дает достоверную, необходимую и полезную ин-



формацию об их структуре и динамике в любом фазовом состоянии, в том числе и аморфном.

Результаты проведенных ранее исследований 2БФМ опубликованы в работах [14–31]. Целью их теоретической части стала проверка используемых методов и приближений применительно к 2БФМ с учетом особенностей его строения в различных полиморфных модификациях и стеклофазе и интерпретация его колебательных спектров, измеренных в широком интервале температур. Основные этапы были связаны с оптимизацией структуры, расчетами электрооптических и механических параметров изолированной молекулы 2БФМ и ее комплексов с водородной связью, моделированием их колебательных спектров.

В [14] при комнатной температуре измерены колебательные спектры кристаллического образца 2БФМ, проведен расчет частот нормальных колебаний и их интенсивностей в ИК спектре свободной молекулы на основе полуклассической теории с использованием метода фрагментов [5], реализованного в комплексе программ LEV-100, дана интерпретация колебательных спектров кристаллического 2БФМ, сделан вывод о наличии в нем водородных связей, предложена вероятная модель конформера молекулы. Результаты оценивались как предварительные. В [15] исследовано влияние водородной связи на структуру 2БФМ. В [16–17] впервые проведено квантово-химическое моделирование строения молекулы 2БФМ методом ТФК и расчет ее колебательных спектров. В [18] методом рентгеноструктурного анализа (РСА) определена структура 2БФМ и на основе анализа колебательных спектров, измеренных в широком интервале температур, изучен процесс стеклования. В [19–24] подробно представлены данные РСА по 2БФМ в триклинной модификации, с использованием метода V3LYP в базисе 6-31G(d) программного пакета GAUSSIAN<sup>03</sup> [1], проведено моделирование структуры и колебательных спектров фрагментов реализующегося в данных условиях Н-комплекса, циклического тетрамера молекул 2БФМ, образованного водородными связями. В [25] опубликованы

результаты экспериментальных исследований структуры и спектров образцов 2БФМ в различных фазовых состояниях, на основании которых уточнена кристаллическая модификация 2БФМ, исследованная методом РСА в [26] – моноклинная нестабильная. Работа [27] по колебательным спектрам 2-, 3-, 4БФМ стимулировала квантово-механические исследования их структуры и спектров [28–31].

В итоге были теоретически обоснованы структура 2БФМ в триклинной модификации и структурообразующая роль Н-связи, интерпретированы колебательные спектры 2БФМ. В цитируемых работах авторов моделирование колебательных спектров базировалось на гармоническом приближении. Расхождение значений вычисленных и измеренных частот колебательных спектров было устранено методом масштабирования [8].

В данной статье представлены результаты расчетов колебательных спектров 2БФМ с учетом механического ангармонизма, которые существенно дополняют полученные в [14–31] и завершают спектроскопическую часть исследований: благодаря им дано полное описание динамики 2БФМ и полностью интерпретированы его колебательные спектры.

#### Моделирование

Структурно-динамическая модель молекулы 2БФМ построена методом ТФП V3LYP в базисе 6-31G(d). В рамках использованного метода гамильтониан Хартри–Фока–Рутана включает в себя дополнительно два функционала, обменный и корреляционный, моделирующие соответствующие взаимодействия. При моделировании обменных взаимодействий использован функционал Беке, содержащий три параметра, которые оценены теоретически на основе результатов расчетов большого числа молекулярных систем. Для учета корреляции электронов использован функционал Ли–Янга–Парра. Решение уравнения Шредингера представляется в виде суперпозиции гауссовых функций. Используемый базис содержит поляризационную функцию. Он включает одну базисную функцию для внутренних оболочек, состав-



ленную из шести гауссовых функций, и набор из двух основных и одной поляризационной базисных функций для внешних электронных оболочек атомов. Первая из двух основных базисных функций содержит в себе три, а вторая – одну гауссову функцию. Дополнительная поляризационная функция состоит из одной гауссовой функции [32].

В результате моделирования минимизирована энергия и оптимизирована геометрия молекулы, рассчитаны ее электрооптические параметры. Силовые постоянные рассчитаны двукратным аналитическим дифференцированием и трех- и четырехкратным численным дифференцированием поверхности потенциальной энергии по нормальным координатам в гармоническом и ангармоническом приближениях соответственно. 72 частоты нормальных колебаний вычислены с точностью до второго порядка стационарной теории возмущений. Вычислены интенсивности в ИК спектре и активности в спектре КР нормальных колебаний, частоты их первых обертонов и комбинационные частоты. По формуле [33], связывающей активности нормальных колебаний с дифференциальными сечениями КР, произведен пересчет сечений КР. При минимальном значении кубической силовой постоянной ( $10 \text{ см}^{-1}$ ) оценивались резонансы Ферми. Теоретически найдены 327 резонансов Ферми и 20 – Дарлингга-Деннисона.

#### Результаты и их обсуждение

Строение и обозначения естественных колебательных координат молекулы 2БФМ указаны на рис. 1. Результаты оптимизации геометрии свободной молекулы в подавляющем большинстве находятся в хорошем согласии с данными рентгеноструктурного анализа для стабильной триклинной модификации 2БФМ. Имеющиеся расхождения определяются влиянием водородной связи и характерны для некоторых структурных элементов группы метанола молекулы, составляющих ядро Н-комплекса. Результаты оптимизации геометрии свободной молекулы и Н-комплекса подробно обсуждены в цитируемых выше работах авторов.

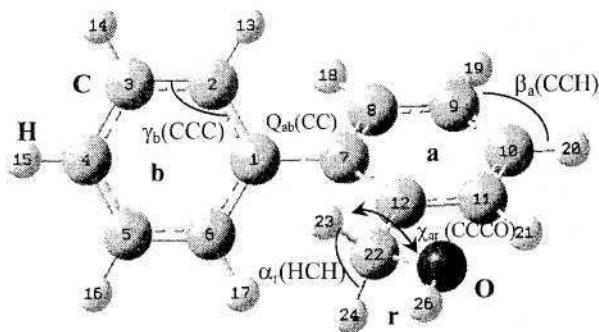


Рис. 1. Строение и естественные колебательные координаты молекулы 2БФМ

В табл. 1 приведены рассчитанные частоты нормальных колебаний молекулы в гармоническом и ангармоническом приближениях и измеренные частоты спектров ИК поглощения и КР и дано их отнесение по рассчитанным коэффициентам форм нормальных колебаний с учетом рассчитанных данных по интенсивностям. Нумерация частот идет в порядке их убывания и соответствует рассчитанным в гармоническом приближении. Как следует из анализа данных, приведенных в табл. 1, учет механического ангармонизма существенно улучшил согласие измеренных и рассчитанных частот валентных колебаний связей СН (область  $3900\text{--}2870 \text{ см}^{-1}$ ): вычисленные частоты уменьшились приблизительно на четыре процента, вплотную приблизившись к экспериментальным. Интерпретация некоторых из них (в таблице они выделены курсивом) уточнилась. Например, частота  $\nu_3$  уменьшилась на  $149 \text{ см}^{-1}$ , а  $\nu_5$  – на  $113 \text{ см}^{-1}$ , став больше  $\nu_3$  на  $28 \text{ см}^{-1}$ . Частота  $\nu_6$  понизилась на  $155 \text{ см}^{-1}$ , больше по сравнению с частотами  $\nu_7$ ,  $\nu_8$ ,  $\nu_9$ , и стала меньше них. В области  $1620\text{--}300 \text{ см}^{-1}$ , в которой проявляются валентные и деформационные колебания фенильных колец и группы метанола, рассчитанные значения частот уменьшились на 1–2%, приблизившись к измеренным. В области неплоских деформационных и крутильных колебаний структурных элементов молекулы, содержащих тяжелые атомы (ниже  $300 \text{ см}^{-1}$ ), учет ангармонизма улучшения не дал: некоторые частоты увеличились по сравнению с рассчитанными в гармоническом приближении,



еще больше удалившись от экспериментальных, а некоторые оказались ниже экспериментальных. Причина этого связана с ограниченностью модели в применении к колебаниям такого типа, для которых характерны широкие потенциальные кривые с неглубокими минимумами и с накоплением систематической ошибки расчета, которое в большей степени сказывается на низких частотах. Критическое отношение к результатам, полученным в низкочастотной области, вынуждает нас в минимальной мере использовать их в колебательном анализе измеренных спектров.

Таблица 1

Измеренные и вычисленные частоты фундаментальных колебаний 2БФМ

№	$\nu_{\text{эксп}}^{\text{ИК}}(\text{I})$ , см <sup>-1</sup>	$\nu_{\text{эксп}}^{\text{КР}}$ , см <sup>-1</sup>	$\nu_{\text{ангарм}}^{\text{г}}$ , см <sup>-1</sup>	$\nu_{\text{гарм}}^{\text{г}}$ , см <sup>-1</sup>	Форма колебания
1	3254	—	3604	3763	$\nu_{\text{R}}(\text{OH})$
2	—	—	3101	3233	$\nu_{\text{A}}(\text{CH})$
3	3062	3062	3061	3211	$\nu_{\text{B}}(\text{CH})$
4	—	—	3073	3206	$\nu_{\text{A}}(\text{CH})$
5	3087	—	3089	3202	$\nu_{\text{B}}(\text{CH})$
6	—	—	3041	3196	$\nu_{\text{B}}(\text{CH})$
7	—	3046	3057	3192	$\nu_{\text{A}}(\text{CH})$
8	—	—	3048	3187	$\nu_{\text{B}}(\text{CH})$
9	—	—	3044	3181	$\nu_{\text{A}}(\text{CH})$
10	3021	—	3023	3180	$\nu_{\text{B}}(\text{CH})$
11	2896	2897	2876	3028	$\nu_{\text{R}}(\text{CH})$
12	2878	2876	2853	2981	$\nu_{\text{R}}(\text{CH})$
13	1605	1615	1621	1662	$\nu_{\text{B}}(\text{CC}), \nu_{\text{B}}(\text{CCH})$
14	1597	1604	1620	1661	$\nu_{\text{A}}(\text{CC}), \nu_{\text{A}}(\text{CCH})$
15	1581	1581	1595	1638	$\nu_{\text{B}}(\text{CC}), \nu_{\text{B}}(\text{CCH})$
16	1568	1568	1588	1627	$\nu_{\text{A}}(\text{CC}), \nu_{\text{A}}(\text{CCH})$
17	—	1501	1515	1550	$\nu_{\text{B}}(\text{CCH}), \nu_{\text{BA}}(\text{CC})$
18	1483	—	1501	1539	$\alpha_{\text{R}}(\text{HCH}), \chi_{\text{AR}}(\text{CCCH})$
19	1475	1479	1494	1527	$\beta_{\text{A}}(\text{CCH}), \alpha_{\text{R}}(\text{HCH})$
20	1463	1464	1469	1499	$\beta_{\text{A}}(\text{CCH}), \beta_{\text{B}}(\text{CCH})$
21	1449	1451	1455	1482	$\beta_{\text{B}}(\text{CCH}), \beta_{\text{A}}(\text{CCH})$
22	1435	—	1431	1470	$\beta_{\text{AR}}(\text{CCH}), \beta_{\text{R}}(\text{COH})$
23	1334	1342	1345	1365	$\nu_{\text{B}}(\text{CCH}), \nu_{\text{B}}(\text{CC})$
24	1313	1305	1319	1348	$\nu_{\text{A}}(\text{CC}), \nu_{\text{A}}(\text{CCH})$
25	1302	1292	1309	1338	$\nu_{\text{B}}(\text{CC}), \nu_{\text{B}}(\text{CCH})$
26	1282	—	1283	1308	$\nu_{\text{BA}}(\text{CC}), \beta_{\text{A}}(\text{CCH})$
27	1278	1282	1280	1305	$\beta_{\text{A}}(\text{CCH}), \nu_{\text{A}}(\text{CC})$
28	—	1226	1242	1277	$\beta_{\text{R}}(\text{HCO}), \beta_{\text{AR}}(\text{CCH})$
29	1218	1209	1221	1242	$\beta_{\text{R}}(\text{COH}), \beta_{\text{A}}(\text{CCH})$
30	1197	1197	1197	1214	$\nu_{\text{B}}(\text{CCH}), \nu_{\text{B}}(\text{CC})$
31	1186	1185	1182	1203	$\beta_{\text{R}}(\text{COH}), \beta_{\text{A}}(\text{CCH})$

Окончание табл. 1

№	$\nu_{\text{эксп}}^{\text{ИК}}(\text{I})$ , см <sup>-1</sup>	$\nu_{\text{эксп}}^{\text{КР}}$ , см <sup>-1</sup>	$\nu_{\text{ангарм}}^{\text{г}}$ , см <sup>-1</sup>	$\nu_{\text{гарм}}^{\text{г}}$ , см <sup>-1</sup>	Форма колебания
32	—	—	1181	1195	$\beta_{\text{A}}(\text{CCH}), \beta_{\text{R}}(\text{COH})$
33	1179	1179	1180	1193	$\nu_{\text{B}}(\text{CCH})$
34	1110	1117 1112	1126	1144	$\beta_{\text{A}}(\text{CCH}), \nu_{\text{R}}(\text{CO})$
35	1078	1088	1087	1109	$\nu_{\text{B}}(\text{CCH}), \nu_{\text{B}}(\text{CC})$
36	1072	1055	1064	1082	$\nu_{\text{A}}(\text{CC}), \beta_{\text{A}}(\text{CCH})$
37	1053	—	1050	1076	$\nu_{\text{R}}(\text{CO})$
38	1035	1034	1046	1065	$\nu_{\text{B}}(\text{CC}), \nu_{\text{B}}(\text{CCH})$
39	1010	—	1020	1044	$\chi_{\text{AR}}(\text{CCCH}), \chi_{\text{R}}(\text{HCOH})$
40	1000	1010	1013	1026	$\gamma_{\text{A}}(\text{CCC}), \gamma_{\text{B}}(\text{CCC})$
41	986	996	1005	1018	$\gamma_{\text{B}}(\text{CCC}), \nu_{\text{B}}(\text{CC})$
42	978	986	983	996	$\chi_{\text{B}}(\text{HCCH}), \chi_{\text{B}}(\text{CCH})$
43	970	977	980	990	$\chi_{\text{A}}(\text{HCCH}), \chi_{\text{A}}(\text{HCCC})$
44	955	951	963	971	$\chi_{\text{B}}(\text{HCCH}), \chi_{\text{B}}(\text{CCH})$
45	946	—	947	956	$\chi_{\text{A}}(\text{HCCH}), \chi_{\text{BA}}(\text{CCCH})$
46	922 917	—	916	933	$\chi_{\text{B}}(\text{HCCH}), \chi_{\text{AB}}(\text{CCCH})$
47	879 874	852	877	892	$\chi_{\text{A}}(\text{HCCH}), \chi_{\text{BA}}(\text{CCCH})$
48	851 841	828	844	865	$\chi_{\text{B}}(\text{CCH}), \chi_{\text{B}}(\text{HCCH})$
49	—	777	802	813	$\nu_{\text{AR}}(\text{CC}), \gamma_{\text{A}}(\text{CCC})$
50	776	762	782	792	$\chi_{\text{A}}(\text{CCCH}), \chi_{\text{B}}(\text{HCCC})$
51	761	754	756	768	$\chi_{\text{A}}(\text{CCCH}), \chi_{\text{B}}(\text{HCCC})$
52	749	745	741	750	$\chi_{\text{A}}(\text{CCCC}), \chi_{\text{A}}(\text{CCCH})$
53	714	—	729	736	$\chi_{\text{A}}(\text{CCCH}), \chi_{\text{A}}(\text{CCCC})$
54	704	714	706	718	$\chi_{\text{B}}(\text{HCCC}), \chi_{\text{B}}(\text{CCCC})$
55	624	625	630	636	$\gamma_{\text{B}}(\text{CCC}), \nu_{\text{B}}(\text{CCH})$
56	616	616	623	631	$\gamma_{\text{B}}(\text{CCC}), \gamma_{\text{A}}(\text{CCC})$
57	605	—	589	593	$\gamma_{\text{AR}}(\text{CCO}), \chi_{\text{A}}(\text{CCCC})$
58	557	559	563	570	$\chi_{\text{B}}(\text{CCCC}), \chi_{\text{B}}(\text{HCCC})$
59	530 520	—	544	549	$\chi_{\text{A}}(\text{CCCC}), \chi_{\text{A}}(\text{HCCC})$
60	491 485	—	458	466	$\chi_{\text{A}}(\text{CCCC}), \chi_{\text{AR}}(\text{CCCH})$
61	430 424	413	430	437	$\chi_{\text{B}}(\text{CCCC}), \chi_{\text{B}}(\text{CCCH})$
62	—	408	410	419	$\chi_{\text{B}}(\text{CCCC}), \chi_{\text{B}}(\text{HCCC})$
63	—	356	337	343	$\chi_{\text{A}}(\text{CCCC}), \chi_{\text{AR}}(\text{CCCC})$
64	—	298	306	310	$\nu_{\text{BA}}(\text{CC}), \gamma_{\text{BA}}(\text{CCC})$
65	—	292	287	282	$\chi_{\text{R}}(\text{HCOH}), \chi_{\text{AR}}(\text{CCOH})$
66	—	376	352	254	$\chi_{\text{AR}}(\text{CCOH}), \chi_{\text{R}}(\text{HCOH})$
67	—	—	235	222	$\chi_{\text{AR}}(\text{CCOH}), \chi_{\text{R}}(\text{HCOH})$
68	—	187	195	176	$\chi_{\text{AR}}(\text{CCCC}), \chi_{\text{AR}}(\text{CCOH})$
69	—	132	105	104	$\chi_{\text{BAR}}(\text{CCCC}), \chi_{\text{BA}}(\text{CCCC})$
70	—	113	82	85	$\chi_{\text{AB}}(\text{CCCC}), \chi_{\text{BAR}}(\text{CCCC})$
71	—	84	72	66	$\chi_{\text{AR}}(\text{CCCH}), \chi_{\text{AR}}(\text{CCCCO})$
72	—	—	47	40	$\chi_{\text{AR}}(\text{CCCH}), \chi_{\text{BA}}(\text{CCCC})$



Получено удовлетворительное согласие рассчитанных интенсивностей нормальных колебаний в спектрах ИК поглощения и сечений КР (рис. 2, 3) с экспериментальными. Исключением стали колебания с участием атома кислорода, что объясняется ограниченностью использованной модели свободной молекулы, не учитывающей образования водородной связи в образце 2БФМ. В первую очередь это относится к валентному колебанию  $\nu(\text{O-H})$ . Влияние водородной связи подробно исследовано в цитируемых работах авторов. Согласно результатам этих исследований, образующийся в кристалле Н-комплекс представляет собой тетрамер молекул 2БФМ симметрии  $C_4$ . Частоты свободной молекулы при его образовании расщепляются на четыре компоненты, две из которых активны в ИКС и две – в СКР. Величины расщеплений в большинстве случаев близки к нулю, и только для частот колебаний струк-

турных элементов Н-комплекса, близких к его ядру, они отличны от нуля. Например, для частоты валентного колебания  $\nu(\text{O-H})$  расщепление максимально и составляет  $141 \text{ см}^{-1}$ ; его определяет дублет, активный в спектре КР ( $3085$  и  $3226 \text{ см}^{-1}$ ), а дублет, активный в ИКС, представлен частотами  $3161$  и  $3203 \text{ см}^{-1}$ . Наличие расщеплений, установленное нами ранее, учитывалось при интерпретации спектров кристаллического 2БФМ на основе результатов расчета спектров молекулы, полученных в ангармоническом приближении.

В табл. 2, 3 приведены измеренные частоты спектров ИК поглощения и КР, которые интерпретируются как обертоны или составные на основании расчета. Эти спектры в значительной мере дополняют друг друга. Только четыре колебательных возбуждения проявляются в обоих спектрах (в таблицах они выделены жирным курсивом):  $1825 \text{ см}^{-1}$

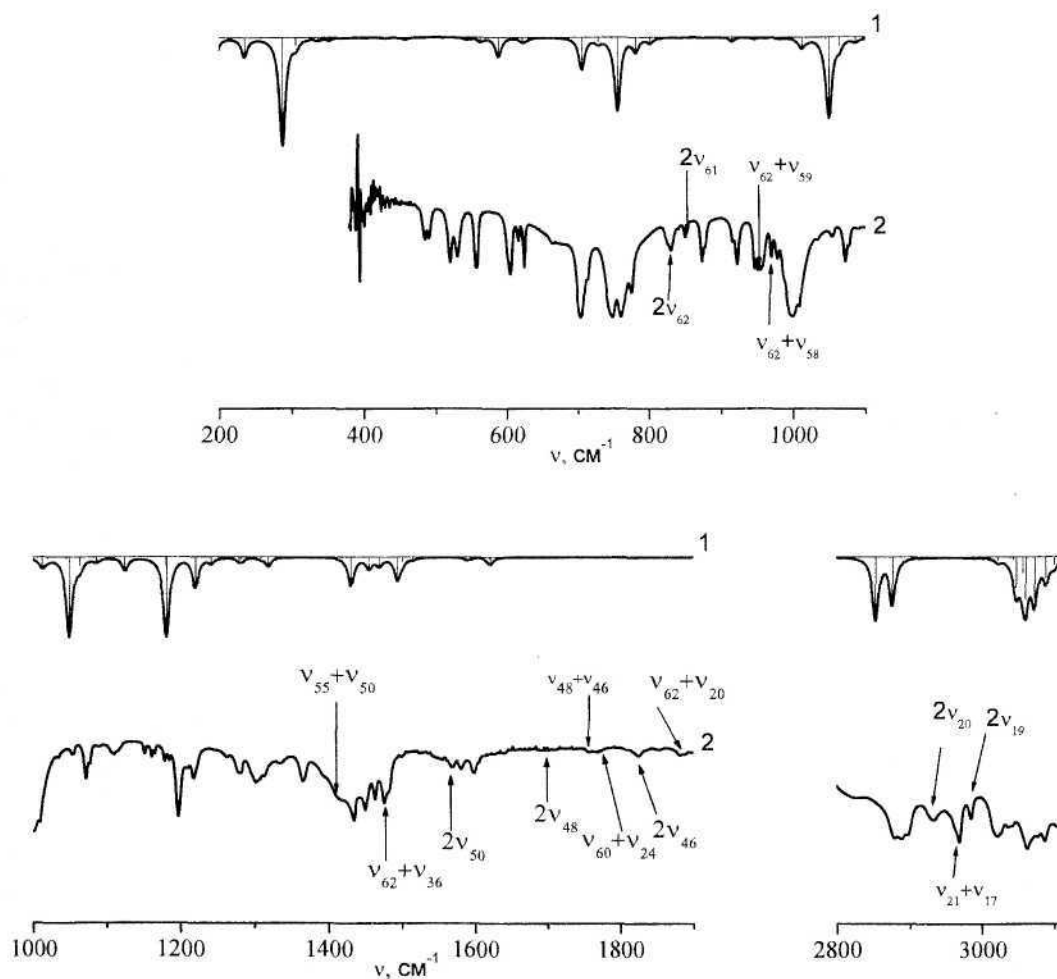


Рис. 2. ИК спектры 2БФМ: 1 – теория, 2 – эксперимент

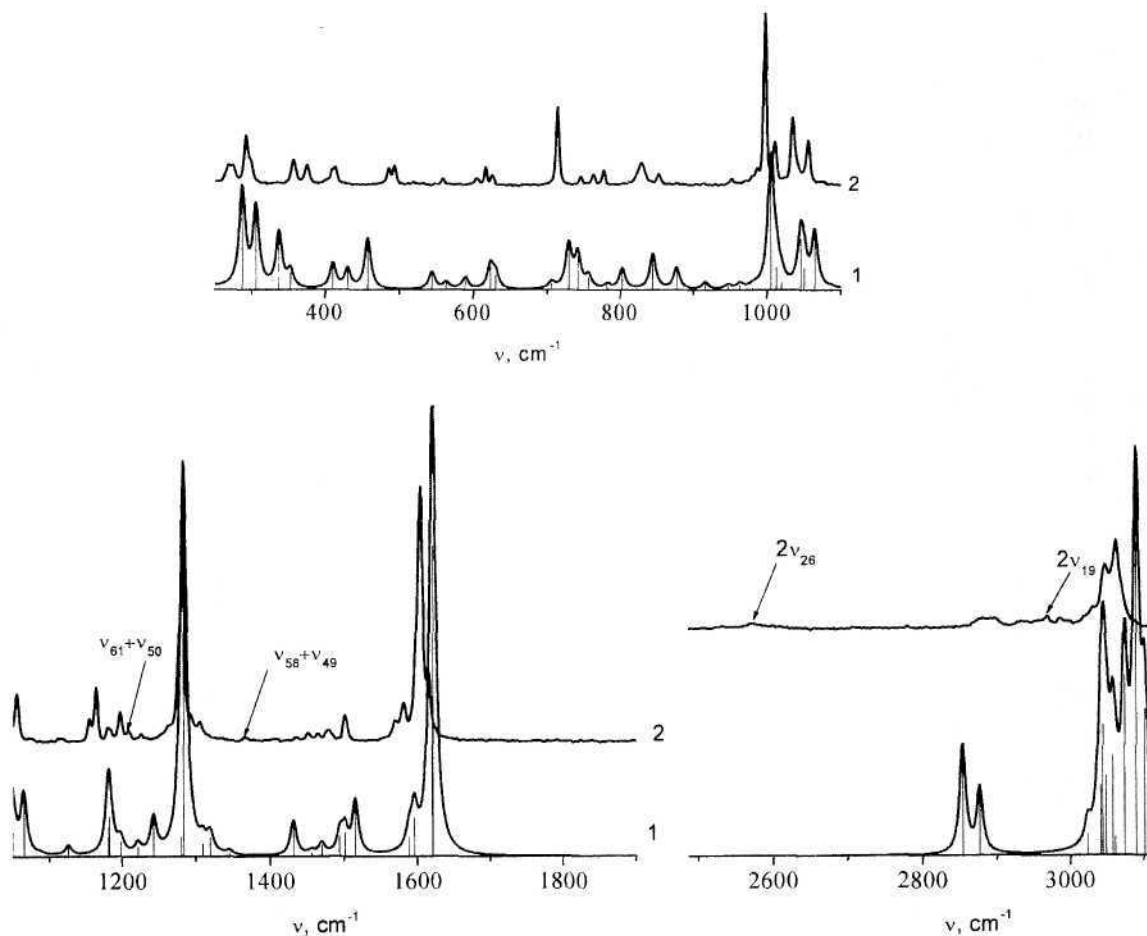


Рис. 3. Спектры КР 2БФМ: 1 – теория, 2 – эксперимент

( $2\nu_{46}$ ),  $1483\text{ см}^{-1}$  ( $2\nu_{52}$ ),  $1210\text{ см}^{-1}$  ( $\nu_{61}+\nu_{50}$ ,  $\nu_{62}+\nu_{49}$ ),  $826\text{ см}^{-1}$  ( $2\nu_{62}$ ). Из-за сильного влияния водородной связи на ИК спектр проявление ангармонизма невозможно наблюдать в области  $3200\text{--}3600\text{ см}^{-1}$ : здесь имеется широкая полоса с аномально большой интенсивностью, соответствующая валентным колебаниям  $q(\text{O-H})$  Н-комплекса. Колебательные возбуждения с обертонами и составными частотами в данной области можно обнаружить только в спектре КР.

Теоретически установленные резонансы Ферми и Дарлинг–Денисона обнаружить в экспериментальных спектрах не удалось, что вполне объяснимо: резонансные взаимодействия невелики, как следует из анализа их матрицы. Ее самые значительные недиагональные элементы не превосходят полпроцента от диагональных. Сказанное проиллюстрируем фрагментом этой матрицы, сим-

метричной относительно главной диагонали, в котором присутствуют самые большие из недиагональных элементов:

$\nu_{32}$	1192	0	-4,12	0,91	0
$\nu_{33}$		1192	0,62	5,42	-3,71
$\nu_{61}+\nu_{52}$			1180	0	0
$\nu_{62}+\nu_{51}$				1171	0
$\nu_{67}+\nu_{44}$					1172

Ее собственные значения и соответствующие им собственные векторы, нормированные на единицу, таковы:

1193	1194	1179	1172	1169
0,95	-0,04	-0,30	0,02	0,04
-0,04	-0,96	0,04	-0,06	0,27
-0,30	-0,30	-0,95	0,02	0
0,03	-0,23	-0,01	-0,04	-0,89
0,01	0,16	-0,02	-0,91	0,37



Таблица 2

Обертоны и составные частоты в ИК спектре 2-БФМ

$\nu_{\text{эксп}}^{\text{ИК}}, \text{см}^{-1}$	$\nu_{\text{ангарм.}}^{\text{ИК}}, \text{см}^{-1}$	$\nu_{\text{гарм.}}, \text{см}^{-1}$	Интерпретация*
2984	2985	3054	2 $\nu_{19}$
2968	2971	3039	$\nu_{20}+\nu_{18}$
2962	2962	3027	$\nu_{20}+\nu_{19}$
2932	2936	2999	2 $\nu_{20}$
1954	1963	1981	2 $\nu_{43}$
1938	1938	1971	$\nu_{51}+\nu_{31}$
1926	1928	1943	2 $\nu_{44}$
1882	1896	1912	2 $\nu_{45}$
<b>1825</b>	<b>1830</b>	<b>1866</b>	2 $\nu_{46}$
1754	1753	1785	2 $\nu_{47}$
1569	1564	1584	2 $\nu_{50}$
<b>1483</b>	<b>1483</b>	<b>1499</b>	2 $\nu_{52}$
1475	1475	1501	$\nu_{62}+\nu_{36}$
1414	1411	1436	2 $\nu_{54}$
1365	1365	1383	$\nu_{58}+\nu_{49}$
1302	1302	1331	$\nu_{60}+\nu_{48}$
1262	1262	1278	$\nu_{70}+\nu_{33}$
<b>1210</b>	<b>1212</b>	<b>1230</b>	$\nu_{61}+\nu_{50}, \nu_{62}+\nu_{49}$
1166	1168	1194	$\nu_{59}+\nu_{56}$
1162	1163	1184	$\nu_{60}+\nu_{54}$
1152	1152	1163	$\nu_{58}+\nu_{57}, \nu_{62}+\nu_{52}$
<b>826</b>	<b>820</b>	<b>838</b>	2 $\nu_{62}$

\* Индексы колебаний в интерпретации соответствуют номерам основных колебаний в табл. 1.

Таблица 3

Обертоны и составные частоты в спектре КР 2-БФМ

$\nu_{\text{эксп}}^{\text{КР}}, \text{см}^{-1}$	$\nu_{\text{ангарм.}}^{\text{КР}}, \text{см}^{-1}$	$\nu_{\text{гарм.}}, \text{см}^{-1}$	Интерпретация*
3576	3586	3746	$\nu_{59}+\nu_6$
3571	3582	3717	$\nu_{53}+\nu_{12}$
3558	3559	3699	$\nu_{54}+\nu_{12}$
3551	3558	3699	$\nu_{60}+\nu_2$
3538	3547	3668	$\nu_{60}+\nu_5$
3416	3417	3552	$\nu_{58}+\nu_{12}$
3206	3211	3287	$\nu_{16}+\nu_{14}$
2968	2971	3039	$\nu_{20}+\nu_{18}$
2881	2886	2952	$\nu_{22}+\nu_{21}$
2875	2876	2943	$\nu_{27}+\nu_{15}$
2872	2874	2935	$\nu_{26}+\nu_{16}$
2867	2871	2932	$\nu_{27}+\nu_{16}$
2844	2846	2904	$\nu_{23}+\nu_{18}$
2814	2814	2939	2 $\nu_{22}$
2779	2779	2838	$\nu_{25}+\nu_{20}$
2722	2722	2781	$\nu_{29}+\nu_{18}$
2681	2686	2730	2 $\nu_{23}$
2619	2620	2677	2 $\nu_{25}$
2513	2514	2583	$\nu_{39}+\nu_{18}, \nu_{40}+\nu_{18}$
2508	2510	2571	$\nu_{46}+\nu_{15}$

Окончание табл. 3

$\nu_{\text{эксп}}^{\text{КР}}, \text{см}^{-1}$	$\nu_{\text{ангарм.}}^{\text{КР}}, \text{см}^{-1}$	$\nu_{\text{гарм.}}, \text{см}^{-1}$	Интерпретация*
2402	2402	2439	$\nu_{45}+\nu_{21}, \nu_{50}+\nu_{14}$
2391	2395	2428	2 $\nu_{30}$
2360	2363	2386	2 $\nu_{33}$
2339	2340	2391	$\nu_{39}+\nu_{24}$
2315	2317	2363	$\nu_{49}+\nu_{17}$
2293	2294	2331	$\nu_{40}+\nu_{27}, \nu_{42}+\nu_{25}$
2267	2267	2304	$\nu_{42}+\nu_{26}, \nu_{35}+\nu_{33}$
2261	2261	2321	$\nu_{39}+\nu_{28}, \nu_{36}+\nu_{30}$
2250	2250	2289	2 $\nu_{34}$
2241	2242	2279	$\nu_{38}+\nu_{30}$
2230	2230	2268	$\nu_{37}+\nu_{33}$
2175			
2170	2173	2219	2 $\nu_{35}$
2128	2129	2151	$\nu_{45}+\nu_{31}, \nu_{45}+\nu_{32}, \nu_{48}+\nu_{26}$
2114	2118	2158	$\nu_{56}+\nu_{19}$
2088	2091	2129	2 $\nu_{38}$
2051	2051	2098	$\nu_{61}+\nu_{14}$
2044	2044	2075	$\nu_{57}+\nu_{21}, \nu_{43}+\nu_{36}$
1984	1984	2008	$\nu_{49}+\nu_{32}$
1879	1880	1919	$\nu_{62}+\nu_{20}$
1873	1873	1901	$\nu_{57}+\nu_{26}$
<b>1824</b>	<b>1830</b>	<b>1866</b>	2 $\nu_{46}$
1604	1604	1625	2 $\nu_{49}$
<b>1478</b>	<b>1483</b>	<b>1499</b>	2 $\nu_{52}$
1451	1458	1472	2 $\nu_{53}$
<b>1209</b>	<b>1212</b>	<b>1230</b>	$\nu_{61}+\nu_{50}, \nu_{62}+\nu_{49}$
1117			
1112	1128	1140	2 $\nu_{58}$
1086	1088	1099	2 $\nu_{59}$
914	915	932	2 $\nu_{60}$
<b>811</b>	<b>820</b>	<b>838</b>	2 $\nu_{62}$

\* Индексы колебаний в интерпретации соответствуют номерам основных колебаний в табл. 1.

В данном случае результаты теории и эксперимента совпали. Как следует из опыта исследований резонансов, если таковые характерны для динамики молекулярной системы, то они явно обнаруживаются в экспериментальных спектрах, а в матрице взаимодействий недиагональные матричные элементы оказываются на порядок выше полученных в нашем случае.

**Выводы**

Учет механического ангармонизма в применении к свободной молекуле 2БФМ существенно улучшил согласие измеренных





и вычисленных частот в высокочастотной и средней областях колебательного спектра. Это дало возможность полнее интерпретировать измеренные колебательные спектры ИК поглощения и КР 2БФМ. Исключение составила частота валентного колебания связи О–Н, которое испытывает значительное влияние водородной связи, не учтенное в модели.

2БФМ, как оказалось, – не вполне удобный объект для моделирования его колебательного спектра в ангармоническом приближении: исследуемый образец есть органический кристалл с водородной связью. Влияние Н-связи на спектр на много сильнее механического ангармонизма в областях, которые соответствуют колебаниям структурных элементов ядра Н-комплекса. Н-комплекс – тетрамер молекул, соединенных Н-связями средней силы. В нем существенно больше колебаний в области ниже  $400\text{ см}^{-1}$ , среди которых колебания, отражающие динамику ядра Н-комплекса. В этой области оба приближения, гармоническое и ангармоническое, работают неудовлетворительно.

#### Список литературы

1. Frisch J., Trucks G.W., Schlegel H.B. Gaussian03, Revision B.03; Gaussian, Inc., Pittsburgh PA, 2003. 302 p.
2. Волькенштейн М.А., Грибов Л.А., Еляшевич М.А., Степанов Б.И. Колебания молекул. М.: Наука, 1972. 700 с.
3. Свердлов Л.М., Ковнер М.А., Крайнов Е.П. Колебательные спектры многоатомных молекул. М.: Наука, 1970. 560 с.
4. Грибов Л.А. Введение в молекулярную спектроскопию. М.: Наука, 1976. 400 с.
5. Грибов Л.А., Дементьев В.А. Моделирование колебательных спектров сложных соединений на ЭВМ. М.: Наука, 1989. 160 с.
6. Pulay P., Fogarasi G., Pongor G., Boggs J.E., Vargha A. Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. Scaled quantum mechanical (QM) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene // J. Amer. Chem. Soc. 1983. Vol.105. P.7037–7047.
7. Yashida H., Takeda K., Okamura J., Ehara A., Matsuura H. A New Approach to Vibrational Analysis of Large Molecules by Density Functional Theory: Wavenumber-Linear Scaling Method // J. Phys. Chem. A. 2002. Vol.106, №14. P.3580–3586.
8. Березин К.В., Кривохижина Т.В., Нечаев В.В. Применение метода линейного масштабирования частот в расчетах нормальных колебаний многоатомных молекул // Оптика и спектроскопия. 2003. Т.94. С.398–401.
9. Willets A., Handy N.C. Anharmonic constants for benzene // Spectrochim. Acta. 1997. Vol.53, №8. P.1169–1177.
10. Березин К.В., Нечаев В.В., Элькин П.М. Ангармонический анализ колебательных состояний пиримидина методом функционала плотности // Оптика и спектроскопия. 2004. Т.97, №2. С.1–11.
11. Березин К.В., Нечаев В.В., Элькин П.М. Ангармонические резонансы в колебательных спектрах пиридазина // Журн. физ. химии. 2005. Т.79, №3. С.1–10.
12. Tanaka H. General view of a liquid-liquid phase transition // Phys. Rev. 2002. Vol.E 62. P.6968–6971.
13. Zallen R. The Physics of Amorphous Solids. N.Y.: Wiley, 1998. 142 p.
14. Babkov L.M., Baran J., Davydova N.A., Kukielski J.I., Trukhachev S.V. Vibrational spectra and structure model of 2-biphenylmethanol molecule // J. Mol. Struct. 2003. Vol.661–662. P.41–48.
15. Babkov L.M., Baran J., Davydova N.A., Trukhachev S.V. The effect of the hydrogen bonds on the molecule conformation in 2-biphenylmethanol // J. Mol. Struct. 2004. Vol.700/1–3. P.55–59.
16. Бабков Л.М., Баран Ж., Давыдова Н.А., Элькин П.М., Кукиельский Ж.И., Успенский К.Е. Квантовомеханическое исследование молекулы 2-би-фенилметанола // Проблемы оптической физики. Саратов, 2004. С.112–115.
17. Babkov L.M., Baran J., Davydova N.A., Elkin P.M., Kukielski J.I., Uspenskiy K.E. Quantum mechanical modeling of the 2-biphenylmethanol molecule // Proc. of SPIE, Saratov Fall Meeting 2003. Laser Physics and Photonics, Spectroscopy and Molecular Modeling IV / Eds. V.L. Derbov, L.A. Melnikov, L.M. Babkov. N.Y., 2004. Vol.5476. P.152–159.
18. Baran J., Davydova N.A., Pietraszko A. Spectroscopic study of the formation of molecular glasses // J. Mol. Struct. 2005. Vol.744–747. P.301–305.
19. Babkov L.M., Baran J., Davydova N.A., Pietraszko A., Uspenskiy K.E. Theoretical and empirical study of 2-biphenylmethanol molecule: the structure and intermolecular interactions // J. Mol. Struct. 2005. Vol.744–747. P.433–438.
20. Бабков Л.М., Баран Ж., Давыдова Н.А., Петрашко А., Успенский К.Е. Исследование структуры и специфических межмолекулярных взаимодействий в 2-бифенилметаноле // Проблемы оптической физики. Саратов: Изд-во Сарат. ун-та, 2005. С.22–26.
21. Uspenskiy K.E., Babkov L.M., Baran J., Davydova N.A., Pietraszko A. Theoretical and empirical investigation of the structure and intermolecular interaction in 2-biphenylmethanol // Proc. of SPIE, Saratov Fall Meeting 2004. Laser Physics and Photonics, Spectroscopy and Molecular Modeling V / Eds. V.L. Derbov, L.A. Melnikov, L.M. Babkov. N.Y., 2006. Vol.5773. P.168–174.
22. Babkov L.M., Baran J., Davydova N.A., Pietraszko A., Uspenskiy K.E. Vibrational spectra and structure modelling of the h-complexes in 2-bi-phenilmethanol using B3LYP/6-31G(d) method // Proc. of SPIE, Saratov Fall Meeting 2005. Laser Physics and Photonics, Spectroscopy and Molecular Modeling / Eds. V.L. Derbov, L.A. Melnikov, L.M. Babkov. N.Y., 2006. Vol.6165. P.616500-1–616500-9.
23. Бабков Л.М., Баран Я., Давыдова Н.А., Петрашко А., Успенский К.Е. Теоретическое и экспериментальные исследования структуры и межмолекулярных взаимодействий в 2-бифенилметаноле // Журн. структ. химии. 2006. Т.47, №4. С.759–766.
24. Babkov L.M., Baran J., Davydova N.A., Uspenskiy K.E. Modeling of the structure and vibrational spectra of 2-biphenylmethanol using B3LYP method // J. Mol. Struct. 2006. Vol.792–793. P.68–72.



25. Baran J., Davydova N.A., Drozd M., Pietraszko A. Experimental evidence of formation of a new phase in supercooled liquid 2-biphenylmethanol // J. Phys.: Condens. Matter. 2006. Vol.18. P.5695–5702.

26. Watkin R., Watkin D. Biphenyl-2-methanol // Acta Cryst. 1995. Vol.C51. P.2161–2163.

27. SDBSWeb: <http://www.aist.go.jp/RIODB/SDBS/> (National Institute of Advanced Industrial Science and Technology. – 29.03.2006).

28. Бабков Л.М., Давыдова Н.А., Баран Я., Успенский К.Е. Колебательные спектры и структурно-динамические модели 2-, 3- и 4-бифенилметанолов // Проблемы оптической физики. Саратов: Изд-во Саратов. ун-та, 2007. С.227–240.

29. Babkov L.M., Baran J., Davydova N.A., Uspenskiy K.E. Vibrational spectra and structural-dynamical models of the 2-,3- and 4-biphenylmethanols // Proc. of SPIE, Saratov Fall

Meeting 2006. Laser Physics and Photonics, Spectroscopy and Molecular Modeling V / Eds. V.L. Derbov, L.A. Melnikov, L.M. Babkov. N.Y., 2007. Vol.6537. P.7–15.

30. Babkov L.M., Baran J., Davydova N.A., Uspenskiy K.E. Investigation of the influence of the hydrogen bonding on the structure and vibrational spectra of biphenylmethanols // Изв. Саратов. ун-та. Нов. сер. Сер. Физика. 2007. Т.7, вып.2. С.44–53.

31. Бабков Л.М., Давыдова Н.А., Успенский К.Е. Водородная связь и структуры 2-, 3- и 4-бифенилметанолов // Журн. структ. химии. 2008. Т.49, №3. С.421–429.

32. Попл Дж.А. Квантово-химические модели // Успехи физ. наук. 2002. Т.172. С.349–356.

33. Polavarapu P.I. Ab Initio vibrational Raman and Raman optical activity spectra // J. Phys. Chem. 1990. Vol.94. P.8106–8112.

УДК 539.193/194:535/33/34

## СТРУКТУРНО-ДИНАМИЧЕСКИЕ МОДЕЛИ ЦИКЛОПЕНТАНА И ЦИКЛОПЕНТЕНА

М.Д. Элькин, Е.А. Джалмухамбетова\*

Саратовский государственный университет  
E-mail: elkinmd@mail.ru

\* Астраханский государственный университет

Проведены модельные квантовые расчеты геометрической структуры и колебательных спектров циклопентана и циклопентена. Установлена некомпланарная геометрическая структура пятичленного цикла соединений.

**Ключевые слова:** колебательные спектры, ангармонические резонансы, конформационные модели циклопентана и циклопентена.

### Structure-Dinamic Modeles of Cyclopropane and Cyclopentene

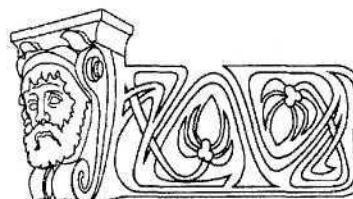
M.D. Elkin, E.A. Djalmuhambetova

Fashionable calculations of geometric and vibrational spectra of cyclopentane and cyclopentene are carried out. Nonplanar conformation for five-member cycle of molecule is cofirmed.

**Key words:** vibrational spectra, anharmonic resonancies, conformational modeles of cyclopentane and cyclopentene.

### Введение

Дискуссия относительно геометрической структуры циклопентана ( $C_5H_{10}$ ) и циклопентена ( $C_5H_8$ ) продолжается в литературе до настоящего времени: является ли пятичленное кольцо плоским? Мнения отечественных и зарубежных спектроскопистов по этому вопросу разделяются.



Аргументы в пользу плоской структуры пятичленного фрагмента, принадлежащего группам симметрии  $D_{5h}$  и  $C_{2v}$  для циклопентана и циклопентена соответственно, приведены в известной монографии [1]. Это и характер поведения ряда полос в колебательных спектрах, и согласие с электронографическими и микроволновыми данными по геометрии молекул.

Теоретическая интерпретация фундаментальных колебательных состояний и оценка электрооптических параметров циклопентана и циклопентена в монографии [1] даны в предположении плоской структуры углеродного кольца. Исходные значения молекулярных параметров были перенесены из парафинов и подвергались варьированию в рамках методики решения обратных спектральных задач. Возможные последствия такого подхода, связанного с рядом обременительных предположений, подробно исследованы в диссертации [2].



В той же монографии [1] достаточно подробно обсуждаются экспериментально наблюдаемые спектроскопические эффекты, свидетельствующие, по мнению авторов, о «небольшой изогнутости углеродного кольца». Нечетко выполняются правила произведения для наблюдаемых частот циклопропана и его дейтерированного аналога  $C_6D_6$ . Имеет место расхождение между вычисленными и экспериментальными данными для значений термодинамических функций циклопентана (в предположении симметрии  $D_{5h}$ ).

Прояснить ситуацию с геометрической структурой молекул можно на основании модельных расчетов параметров адиабатического потенциала, поскольку конформационные свойства отдельных молекулярных фрагментов наглядно проявляются в колебательных спектрах соединений. Такие возможности предоставляют неэмпирические квантовые методы расчета электронной структуры соединений в адиабатическом приближении [3].

Цель данной работы – построение информационных структурно-динамических моделей циклопентана и циклопентена с использованием метода функционала плотности DFT/b3LYP [4] с различными базисными наборами [5].

#### Анализ моделей

Оптимизация геометрии осуществлялась в рамках конформационных моделей различной исходной симметрии. При этом для пятичленного углеродного фрагмента предполагалась симметрия  $D_{5h}$ , а симметрия исследуемого соединения задавалась положением фрагментов  $CH_2$  относительно плоскости пятичленного кольца.

Для исходных конфигураций циклопентана и циклопентена, принадлежащих группам симметрии  $D_{5h}$  и  $C_{2v}$  соответственно, симметрия оптимизированной геометрии сохранялась, но в расчетах спектра фундаментальных колебаний появлялись отрицательные значения, по абсолютной величине превышающие  $200\text{ см}^{-1}$ . Ошибками численных методов технологии «Gaussian» это не объяснить [5], а тот факт, что это касается неплоских колебаний углеродного кольца, по-

зволяет сделать вывод о некомпланарности атомов этого кольца.

Все остальные исходные модели, отличающиеся положениями фрагментов  $CH_2$  относительно плоского пятичленного цикла, симметрией не обладали, а в процессе оптимизации геометрии приводили к конфигурации молекул с симметрией, близкой к группе  $C_2$  с точностью до третьего знака мантиссы в значениях координат подобных атомов. Этот факт указывает на наличие глобального минимума адиабатического потенциала, соответствующего рассчитанной геометрической структуре. Для такой неплоской конфигурации атомов углеродного пятичленного фрагмента хорошо воспроизводятся частоты фундаментальных колебаний в низкочастотной области спектра, интерпретируемые как колебания типа  $\rho$  и  $\chi$  – изменение углов между связями  $CH$  и  $CC$  соответственно и плоскостями  $CCC$  пятичленного углеродного кольца. Наглядно это иллюстрируется в табл. 1, 2. Отметим, что полученные результаты согласуются с теорией «псевдповращения» Питцера [1], согласно которой атомы углерода пятичленного фрагмента периодически последовательно выходят из плоскости кольца.

Степень изогнутости углеродного кольца, о которой велась дискуссия в монографии [1], можно оценить из результатов проведенной оптимизации геометрии для конфигурации молекул циклопентана и циклопентена с симметрией  $C_2$ .

Для циклопентана значения длин валентных связей равны:  $R_{CC} = 1.54\text{--}1.56\text{ \AA}$ ,  $R_{CH} = 1.09\text{--}1.1\text{ \AA}$ ; валентные углы углеродного кольца  $A_{CCC} = 103.3\text{--}106.1^\circ$ ; двугранные углы углеродного кольца  $D(1,2,3,4) = 12.8^\circ$ ,  $D(2,1,5,4) = 41.5^\circ$ ,  $D(3,4,5,1) = -33.5^\circ$ . Отступление от тетраэдричности для углов  $A_{CCN}$  и  $A_{HCH}$  не превышает  $3.8^\circ$ .

Эти данные хорошо согласуются с электронографическим экспериментом [1], учитывая математические модели теории для оценки геометрических параметров такого эксперимента [6–8].

Для циклопентена –  $R_{C1C2} = 1.51\text{ \AA}$ ,  $R_{C2C3} = 1.54\text{ \AA}$ ,  $R_{C1-C5} = 1.33\text{ \AA}$ ,  $R_{CH} = 1.09\text{--}1.1\text{ \AA}$ ; валентные углы углеродного кольца  $A(2,1,5) = 112.1^\circ$ ,  $A(1,2,3) = 103.1^\circ$ ,  $A(2,3,4) =$



Таблица 1

Влияние базиса на интерпретацию колебательного спектра циклопентана

Форма колебаний	Циклопентан – C <sub>5</sub> H <sub>10</sub>									C <sub>5</sub> D <sub>10</sub>	
	ν <sub>exp</sub> [1]	ν <sub>b</sub>	ν <sub>anh</sub>	ИК	КР	ν <sub>b</sub>	ν <sub>anh</sub>	ИК	КР	ν <sub>exp</sub> [1]	ν <sub>m</sub>
		Минимум				Максимум					
	Тип симметрии А										
qcn	2966	3075	2927	18.9	119	3099	2949	25.3	123	2214	2227
qcn	2966	3071	2936	11.9	184	3093	2950	16.2	277	2214	2203
qcn	2966	3045	2932	54.8	122	3066	2938	59.2	143	2125	2156
qcn	2878	3027	2866	0.01	73.9	3049	2966	0.27	95.1	2125	2135
qcn	2878	3009	2868	17.1	140	3036	2906	20.4	198	2125	2112
ν <sub>ncn</sub>	1462	1524	1481	0.03	0.80	1553	1520	0.16	9.81	1155	1184
ν <sub>ncn</sub>	1453	1499	1461	0.94	1.88	1527	1490	3.83	4.47	1155	1154
ν <sub>ncn</sub>	1453	1488	1454	0.08	13.2	1516	1480	0.21	26.6	1124	1125
ν <sub>scn</sub>	1316	1350	1315	0.02	0.17	1371	1339	0.79	0.57	1064	1086
ν <sub>scn</sub>	1295	1339	1301	0.03	0.87	1362	1322	0.08	2.45	1050	1072
ν <sub>scn</sub>	1243	1283	1262	0.87	7.63	1297	1275	0.94	17.8	1021	1014
ν <sub>scn</sub>	1207	1236	1199	0.10	0.41	1252	1212	0.24	1.27	937	935
ν <sub>scn</sub>	–	1180	1147	0.00	4.11	1194	1163	0.01	10.8	874	898
ν <sub>scn</sub>	1020	1040	1013	0.00	6.91	1048	1023	0.00	8.26	778	792
ν <sub>scn</sub>	1004	1003	974	0.42	0.93	1013	988	0.63	2.04	778	771
Q <sub>cc</sub>	896	907	883	1.39	4.09	916	895	2.43	5.31	738	742
Q <sub>cc</sub>	886	881	863	0.35	19.8	889	872	0.57	23.6	–	718
ν <sub>scn</sub>	–	822	840	0.12	0.26	829	845	0.20	0.62	–	674
γ <sub>ccc</sub>	628	620	618	0.07	0.51	621	619	0.17	1.24	–	508
γ <sub>ccc</sub>	283	261	259	0.00	0.06	265	264	0.01	0.22	–	204
Тип симметрии В											
qcn	2966	3090	2942	110	48.7	3114	2961	115	53.7	2220	2232
qcn	2966	3073	2927	31.7	99.1	3097	2949	33.5	111	2220	2216
qcn	2966	3059	2919	38.2	5.56	3084	2935	41.8	7.77	–	2186
qcn	2878	3027	2874	51.2	25.1	3050	2894	55.1	32.1	2132	2143
qcn	2878	3016	2871	38.7	13.6	3041	2905	40.1	17.7	2132	2136
ν <sub>ncn</sub>	1462	1499	1463	2.23	2.00	1527	1490	4.90	3.90	–	1087
ν <sub>ncn</sub>	1453	1494	1456	3.21	9.94	1522	1485	4.85	18.7	1064	1075
ν <sub>scn</sub>	1312	1349	1313	0.05	0.12	1371	1334	0.81	0.74	1021	1038
ν <sub>scn</sub>	1283	1321	1277	0.51	4.80	1341	1291	0.77	14.5	–	980
ν <sub>scn</sub>	1283	1310	1276	0.00	0.03	1323	1289	0.00	0.08	943	934
ν <sub>scn</sub>	1207	1227	1188	0.40	4.77	1243	1216	0.60	13.1	802	800
ν <sub>scn</sub>	–	1206	1180	0.22	1.03	1220	1199	0.32	2.39	738	751
ν <sub>scn</sub>	1020	1048	1012	0.02	5.79	1060	1044	0.04	7.51	738	726
Q <sub>cc</sub>	–	973	964	0.96	0.22	981	973	1.26	0.32	1155	1159
Q <sub>cc</sub>	886	902	878	1.36	0.16	910	887	2.52	0.26	1124	1121
ν <sub>scn</sub>	–	868	842	0.50	0.35	876	875	0.76	1.49	–	696
ν <sub>scn</sub>	717	768	782	0.70	0.00	776	788	0.89	0.08	570	563
γ <sub>ccc</sub>	546	553	553	1.03	0.30	553	554	1.24	0.46	438	424
xx	–	36	30	0.00	0.01	40	50	0.01	0.02	–	27

Примечание. Частоты колебаний ν в см<sup>-1</sup>, интенсивности в спектрах ИК в Км/моль, в спектрах КР – Å<sup>4</sup>/а.е.м.



Таблица 2

## Влияние базиса на интерпретацию колебательных состояний цикlopентана

Форма колебаний	$\nu_{\text{exp}} [1]$	$\nu_h$	$\nu_{\text{anh}}$	ИК	КР	$\nu_h$	$\nu_{\text{anh}}$	ИК	КР
		Минимум				Максимум			
Тип симметрии $A'$									
qcn	3066o.c	3188	3052	22.6	161	3213	3077	32.4	188
qcn	2958o.c	3090	2941	43.1	77.8	3113	2962	55.6	89.6
qcn	2927o.c	3055	2922	55.1	125	3074	2970	67.7	171
qcn	2916	3035	2906	5.25	53.6	3055	2966	8.54	67.3
qcn	2852o.c	2996	2792	31.2	185	3015	2911	39.2	220
Qc-c	1616	1671	1632	1.64	14.1	1702	1663	4.30	34.3
$\beta_{\text{HC}}$	1468	1508	1470	0.62	5.71	1539	1498	2.79	19.1
$\beta_{\text{HC}}$	1449	1483	1445	0.70	18.1	1513	1475	1.71	28.1
$\beta_{\text{CC}}$	1302	1325	1294	0.05	2.69	1347	1317	1.34	3.23
$\beta_{\text{CC}}$	1207ep	1229	1201	1.24	6.29	1243	1216	2.00	13.3
$\beta_{\text{CC}}$	1108	1129	1117	0.08	18.5	1140	1127	0.19	22.9
$\beta_{\text{CC}}$	1048o.c	1068	1046	4.81	0.77	1083	1061	8.49	1.58
$\beta_{\text{CC}}$	965	972	956	0.59	5.82	981	966	1.33	8.34
Qcc	900	902	883	0.35	11.2	909	891	0.62	14.6
Qcc	800	820	799	0.46	1.06	828	807	1.03	2.17
pcn	692o.c	710	700	24.4	1.58	718	707	40.9	9.09
$\gamma_{\text{CC}}$	603	611	612	10.0	0.91	618	616	15.2	3.57
xx	—	139	117	0.19	0.69	144	135	0.29	0.81
Тип симметрии $A''$									
qcn	3060c	3163	3048	7.84	82.1	3189	3073	12.5	92.2
qcn	2902o.c	3043	2903	32.1	116	3062	2922	43.3	129
qcn	2852o.c	2997	2814	53.914	59.0	3015	2867	62.4	77.9
$\beta_{\text{HC}}$	1468	1488	1446	0.22	10.3	1520	1478	2.93	14.5
$\beta_{\text{CC}}$	1356	1377	1348	0.69	0.28	1398	1368	2.19	0.52
$\beta_{\text{CC}}$	1297	1317	1287	0.01	1.26	1336	1304	0.45	5.72
$\beta_{\text{CC}}$	1284	1307	1280	1.69	0.70	1323	1296	2.64	1.70
$\beta_{\text{CC}}$	1207	1224	1195	0.00	4.92	1240	1211	0.06	12.1
$\beta_{\text{CC}}$	1113	1149	1128	1.30	0.39	1165	1145	1.64	0.75
$\beta_{\text{CC}}$	1027	1034	1010	1.52	4.59	1046	1024	2.05	5.94
$\beta_{\text{CC}}$	965	964	948	0.06	0.11	974	956	0.13	2.33
Qcc	904o.c	916	901	5.19	0.18	927	912	7.77	0.32
Qcc	—	889	873	2.51	1.23	901	886	4.82	1.56
$\gamma_{\text{CC}}$	772сл	778	766	0.18	0.82	785	773	0.38	1.19
pcn	387	390	381	0.05	1.82	393	393	0.07	4.06

= 105.6°; двугранные углы углеродного кольца  $D(6,1,5,13) = 0$ ,  $D(1,2,3,4) = 19.1^\circ$ ,  $D(3,4,5,1) = 12.4^\circ$ . Отступление от тетраэдричности для углов фрагмента  $\text{CH}_2$  не превышает 2.8°. Положение связи  $\text{C}_1\text{H}_7$  относительно кольца определяет двугранный угол  $D(6,1,2,3) = 169.9^\circ$ .

Смена базиса расчета от 6-31G(p) до 6-311++G(d,p) на расчетных значениях гео-

метрических параметров сказывается на третьем разряде мантиссы для валентных связей и на втором разряде для валентных и двугранных углов, что не выходит за рамки точности электронографического или микроволнового эксперимента по определению геометрических параметров молекул.

Влияние выбора базиса расчета на теоретическую интерпретацию спектра фунда-



ментальных колебаний циклопентана и циклопентена продемонстрировано в табл. 1, 2.

Располагая экспериментальными данными по колебательным спектрам дейтерозамещенного циклопентана, можно проверить надежность полученной для циклопентана  $C_5H_5$  системы гармонических силовых постоянных, решая прямую механическую задачу для молекулы  $C_5D_{10}$ , используя поле базовой молекулы. Результаты представлены в последних колонках табл. 1. После операции масштабирования по соотношению  $v_m = v_h(0.9813 - 0.835 \times 10^{-6} v_h)$  [9] имеем хорошее совпадение экспериментальных и теоретически рассчитанных частот фундаментальных колебаний для циклопентана-D10.

Данные расчета фундаментальных состояний циклопентана и циклопентена в гармоническом приближении теории молекулярных колебаний (столбец  $v_h$  в табл. 1, 2) указывает на существенное расхождение с экспериментом для высокочастотного диапазона. Естественно, в первую очередь это касается валентных колебаний связей СН. Причина известна – не учтен ангармонизм колебаний.

Существуют два подхода к решению этой проблемы. Первый подход базируется на использовании операции масштабирования как методе учета погрешностей квантовых расчетов гармонических силовых полей. Вводимые масштабирующие множители в линейном или квадратичном соотношении подбирались на основании статистической обработки результатов теоретического анализа колебательного спектра, его сопоставления с оптическим экспериментом.

Существуют различные схемы масштабирования расчетных значений фундаментальных колебаний и системы гармонических силовых полей исследуемых молекулярных соединений. Подробно эти вопросы освещены в диссертации [9] и публикации [10].

При таком подходе возникает естественный вопрос: является ли операция масштабирования компенсацией ненадлежащего учета электронной корреляции в расчетных методах квантовой механики молекул (квантовой химии) или это метод эмпирического учета ангармонизма колебаний? Ответ на

этот вопрос, на наш взгляд, можно получить, осуществляя расчеты параметров адиабатического потенциала исследуемого соединения в ангармоническом приближении.

Такие расчеты для ряда соединений различных классов представлены в публикациях [11–16]. На их основе сделаны выводы, что метод функционала плотности [4], в определенной мере учитывающий электронную корреляцию [6], позволяет оценить систему гармонических силовых постоянных из квантовых расчетов, а операцию масштабирования следует рассматривать как метод учета ангармонизма колебаний.

Для учета влияния ангармонизма воспользуемся модельным гамильтонианом из оригинальной монографии [17]:

$$H^{(v)} = \frac{1}{2} \{ (P_s^2 + (v_s Q^s)^2 + F_{sri} Q^s Q^r Q^i + F_{stru} Q^s Q^r Q^i Q^u \}. \quad (1)$$

Здесь  $v_s$  – частоты гармонических колебаний;  $Q^a$  – нормальные колебательные координаты, линейно связанные с декартовыми координатами ядер;  $F_{sri}$  и  $F_{stru}$  – кубические и квартичные силовые постоянные.

Решение уравнения (1) методами теории возмущения приводит к известному выражению для энергий колебательных состояний:

$$E^{(v)} = v_s(n_s + g_s/2) + \chi_{sr}(n_s + 1/2)(n_r + 1/2)(1 + 1/2 \delta_{sr}), \quad (2)$$

где  $n_s$  – квантовые числа колебательного состояния,  $g_s$  – степень вырождения колебательных координат,  $\chi_{sr}$  – ангармонические постоянные.

Выражения для ангармонических силовых постоянных, полученных во втором порядке адиабатической теории возмущения, имеют вид:

$$\begin{aligned} \chi_{ii} &= \frac{1}{4} \{ 6F_{iii} - 15(F_{ii})^2/v_i - \Sigma(F_{ij})^2 / v_j(8v_i^2 - 3v_j^2)/(4v_i^2 - v_j^2) \}, \\ \chi_{ij} &= \frac{1}{4} \{ F_{ijj} - 6(F_{ij}F_{ij}/v_i) - 4F_{ij}^2[v_i/(4v_i^2 - v_j^2)] - \Sigma(F_{ik} - (F_{kj})^2/v_k) - F_{ijk}v_k(v_k^2 - v_i^2 - v_j^2) / [(v_i + v_k + v_k)(v_i + v_k - v_k)(v_i - v_k + v_k)(v_i - v_k - v_k)] \}. \end{aligned} \quad (3)$$

Из соотношений (3) нетрудно получить ангармонические поправки для сдвига фун-



даментальных состояний при учете ангармонизма колебаний:  $X_{ii} = 2 \chi_{ii}$  и  $X_i = 1/2 \sum \chi_{ir}$ .

Резонансные частотные функции (знаменатели) в выражениях (3) ограничивают применение этих соотношений в случае малой энергетической щели между фундаментальным состоянием  $v_i$  и обертоном  $v_j + v_k$ .

В технологии «Gaussian-03», по умолчанию, условием применимости соотношения (3) является значение щели  $v_i - v_j - v_k$  более  $10 \text{ см}^{-1}$ . В противном случае схема (3) в алгоритме, заложенном в технологии «Gaussian-03», заменится на схему учета резонансов Ферми, предложенную в монографии [18]:

$$E = \{(v_i + v_j + v_k) \pm (F_{ijk})^2 (1 + \delta_{jk})/2 + ((v_i - v_j - v_k)^2)^{0.5}\}/2. \quad (4)$$

Условие, налагаемое на применимость соотношений (3) и (4), в каждом численном эксперименте надо рассматривать индивидуально, поскольку это общее условие применимости теории возмущения. В технологии «Gaussian-03» предусмотрена возможность установления нужного значения энергетической щели для резонирующих колебательных состояний.

Использование технологии «Gaussian-03» с параметрами по умолчанию является основной причиной того, что ангармонические смещения фундаментальных полос для разных базисных наборов могут отличаться. Нередко на этом основании делается вывод о преимуществе того или иного базиса, хотя значения соответствующих кубических и квартичных силовых постоянных, входящих в выражения (3), от выбора базиса существенно не зависят [11–16]. Изменение порогового значения щели может снять остроту вопроса о «лучшем базисном наборе» при анализе колебательных состояний конкретного молекулярного соединения.

Для молекул циклопентана и циклопентена число резонирующих колебательных состояний при щели в  $10 \text{ см}^{-1}$  меняется в пределах от 12 до 23. Расхождение в вычисленных значениях фундаментальных состояний для частот валентных колебаний связей СН, как это следует из табл. 1, 2, может достигать  $100 \text{ см}^{-1}$  в зависимости от выбранного базиса. Лучшее согласие с экспериментом

дают базисы 6-311G(d,p) и 6-311+G(d,p). При этом влиянием поляризационных (p) и диффузионных эффектов (+, ++) можно пренебречь.

Следует подчеркнуть, что эксперимент по колебательным спектрам циклопентана и циклопентена, приведенный в монографии [1], касается жидкой фазы вещества, поэтому полосы имеют плохое разрешение, растянутые максимумы, размыты. В большей степени это относится к полосам, интерпретированным как валентные колебания связей СН. Для остальных полос выбор базиса расчета принципиального значения для интерпретации колебательных состояний не имеет.

Нелишне подчеркнуть, что численный эксперимент проведен для свободной молекулы. Сравнение таких расчетов корректнее проводить для газообразной фазы вещества и его паров. Тем не менее имеет место качественное согласие вычисленных и экспериментальных значений интенсивностей колебательных полос. Наиболее интенсивными являются полосы, интерпретированные как колебания связей СН. Характер поведения полос в спектрах ИК и КР в области ниже  $1500 \text{ см}^{-1}$ , который послужил аргументом, приведенным в монографии [1], при доказательстве плоской структуры углеродного кольца, сохраняется для неплоской модели пятичленного цикла.

## Выводы

1. Неэмпирические квантовые расчеты параметров адиабатического потенциала позволяют интерпретировать колебательные состояния циклопентана и циклопентена только в предположении неплоской конформации углеродного кольца.

2. Для получения хорошего согласия с экспериментом по колебательным состояниям следует производить расчеты в ангармоническом приближении теории молекулярных колебаний.

3. Квантовые расчеты позволяют оценить гармоническое силовое поле исследуемых молекул. Операцию масштабирования следует рассматривать как метод оценки влияния ангармонизма на смещение фундаментальных полос.



## Список литературы

1. Свердлов Л.М., Ковнер М.А., Крайнов Е.П. Колебательные спектры многоатомных молекул. М.: Наука, 1970. 560 с.
2. Березин В.И. Прямые и обратные задачи спектроскопии циклических и комплексных соединений: Дис. ... д-ра физ.-мат. наук. Саратов, 1983. 336 с.
3. Минкин В.И. Теория строения молекул. Ростов н/Д: Феникс, 1997. 386 с.
4. Yoshida H., Takeda K., Okamura J. et al. New Approach to Vibrational Analysis of Large Molecules by Density Functional Theory: Wavenumber-Linear Scaling Method // J. Phys. Chem. A. 2002. Vol.106, №14. P.3580–3586.
5. Frisch M.J., Trucks G.W., Schlegel H.B. et al. Revision A.7. Gaussian, Inc., Pittsburgh (PA), 2003.
6. Элькин М.Д. Кинематическая ангармоничность в электронно-графических исследованиях геометрии молекул // Журн. структ. химии. 1986. Т.27. С.42–46.
7. Элькин М.Д. К вопросу об использовании функции плотности вероятности в ангармонической теории рассеяния электронов молекулами // Журн. структ. химии. 1989. Т.30, №6. С.33–37.
8. Элькин М.Д., Костерина Э.К. Внутримолекулярная динамика и её математическое описание в задачах молекулярной спектроскопии и газовой электронографии // Хим. физика. 1994. Т.10, №1. С.38–42.
9. Березин К.В. Квантово-механические модели и решение на их основе прямых и обратных спектральных задач для многоатомных молекул: Дис. ... д-ра физ.-мат. наук. Саратов, 2004. 264 с.
10. Краснощеков С.В., Степанов Н.Ф. Масштабирующие множители как эффективные параметры для коррекции неэмпирического силового поля // Журн. физ. химии. 2007. Т.81, №4. С.680–689.
11. Элькин М.Д., Эрман М.А., Пулин В.Ф. Структурно-динамические модели и ангармонический анализ колебательных состояний пятичленных циклических соединений // Вестн. Саратов. гос. техн. ун-та. 2006. №4, вып.4. С.38–44.
12. Элькин П.М., Эрман М.А. Структурно-динамические модели и ангармонический анализ колебательных состояний полихлорзамещенных дибензо-*p*-диоксинов // Журн. прикл. спектроскопии. 2007. Т.74, №1. С.21–24.
13. Элькин М.Д., Эрман Е.А., Пулин В.Ф. Колебательные спектры конформеров бензофенона // Журн. прикл. спектроскопии. 2007. Т.74, №5. С.563–568.
14. Элькин М.Д., Джалмухамбетова Е.А., Гречухина О.Н. Проявление межмолекулярного взаимодействия в димерах урацила // Изв. Саратов. ун-та. Нов. сер. 2008. Сер. Физика. Т.8, №2. С.24–30.
15. Элькин П.М., Пулин О.В., Джалмухамбетова Е.А. Теоретический анализ таутомерных форм пурина // Журн. прикл. спектроскопии. 2008. Т.75, №1. С.23–27.
16. Элькин П.М., Эрман Е.А., Пулин О.В. Квантово-химический расчет нормальных колебаний молекул замещенных пятичленных халькоген-гетероциклических соединений с учетом ангармонизма анализ структуры и спектров пятичленных циклических соединений // Журн. прикл. спектроскопии. 2009. Т.76, №2. С.170–175.
17. Amat G., Nielsen H.H., Torrago G. Rotation-vibration of polyatomic molecules. N.Y.: Pergamon Press, 1971. 580 с.
18. Герцберге Г. Электронные спектры и строение многоатомных молекул. М.: Мир, 1969. 772 с.

УДК 530.182

## КРИПТОГРАФИЯ ГЛАЗАМИ ФИЗИКА

Ю.Н. Зайко

ФГОУ ВПО Поволжская академия государственной службы им. П.А. Столыпина  
E-mail: zyrnick@rambler.ru

Приведены результаты исследования криптоалгоритмов DES (США) и ГОСТ 28147-89 (Россия) методами нелинейной динамики. Исследуются точечные отображения, задаваемые важными элементами криптоалгоритмов – блоками подстановок (S-блоками). Продемонстрировано явление возврата. Исследована эргодичность рассматриваемых отображений. Оценка качества S-блока может быть выполнена с помощью отображения первого возвращения. Приведены результаты статистического исследования S-блоков. Показано, что отбор S-блоков ГОСТ из всего множества подстановок не связан с их случайным и равновероятным выбором из указанного множества.

**Ключевые слова:** криптография, подстановка, отображение, цикл, нелинейная динамика, явление возврата, эргодичность, ключ, S-блок, динамический хаос, лавинный эффект, перестановка.

## Cryptography from the Physicist's Point of View

Yu.N. Zayko

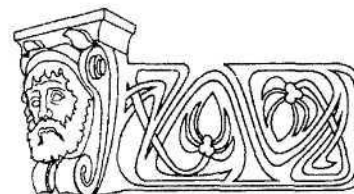
The results of treating of crypto algorithms such as DES (USA) and GOST 28147-89 (Russia) with the help of nonlinear dynamics methods are presented. Point maps which are generated by the blocks

of substitutions (S-blocks) are investigated. The phenomenon of return is demonstrated. The ergodicity of these maps are treated. An estimation of quality of S-block could be made by maps of first return. The results of statistical treating of S-blocks are presented.

**Key words:** cryptography, substitution, mapping, cycle, nonlinear dynamics, return, ergodicity, key, S-box, dynamical chaos, avalanche effect, permutation.

## Введение

Криптография является одним из основных средств защиты информации. Зашифровать информацию можно легко и, самое главное, быстро – сейчас для этого могут быть использованы различные алгоритмы как в программном, так и в аппаратном исполнении (о некоторых оговорках сказано ниже). Выполнить обратное преобразование без знания секретной информации – ключа,







т.е. взломать шифр при современном состоянии криптоанализа, значительно сложнее. Например, для взлома блочного симметричного криптоалгоритма DES (Data Encryption Standard, США, длина ключа 56 бит) методом распределенных вычислений по данным конкурса компании RSA Data Security, проведенного в 1997 г., потребовалось 140 дней, а для криптоалгоритма RC5 с той же длиной ключа – 210 дней. С деталями подобных атак можно познакомиться в [1, 2]. Там же приведены результаты и прогнозы аналогичных атак на асимметричные криптосистемы, требующие значительно больших временных и вычислительных ресурсов.

Если оценивать степень защиты информации по отношению (стоимость взлома)/ (стоимость защиты), то криптографические методы далеко превосходят все прочие.

Правительства всех стран давно используют криптографию для защиты дипломатической и военной информации. Что же касается коммерческой, и, особенно, частной информации, то здесь правительства проявляют сдержанность, предлагая для использования криптосистемы с явно заниженными криптографическими свойствами, например понижая длину ключа. Есть и другие способы затормозить широкое использование криптографии. Билль Сената США № 266, требующий предоставления правительству права получать простое текстовое содержание разговора, данных и других видов связи и Указ Президента РФ № 334 [3] о необходимости использования только сертифицированных криптографических средств, по сути дела, являются только разными по форме мерами установления государственного контроля за использованием криптографии в целях защиты конфиденциальной, и в том числе личной, информации.

#### 1. Сертифицированные криптографические средства

В США к таким средствам относился уже упоминавшийся выше криптоалгоритм DES, разработанный в IBM под руководством Х. Фейстеля под названием Lucifer в 1976 г. и опубликованный в 1977 г. [4]<sup>1</sup>, а в

<sup>1</sup> В 2000 г. ему на смену пришел AES (Advanced Encryption Standard) [6].

России – криптоалгоритм ГОСТ 28147 – 89, разработанный в 1989 г. [5]. Они достаточно подробно описаны, например в [1, 2]. Оба являются симметричными блочными шифрами, т.е. совершают криптографические преобразования с блоками текста длиной 64 бита и имеют ключи длиной 64 (эффективно 56) и 256 бит соответственно. Оба используют схему преобразования, разработанную Фейстелем, заключающуюся в циклической перестановке правой и левой частей блока с одновременным побитовым сложением по модулю 2 очередной правой части и текущего ключа. Отличие только в числе циклов: у DES 16 циклов, а у ГОСТ 32 цикла. Другие детали преобразований мы опустим (они подробно описаны в [1, 2]). Остановимся только на преобразованиях 32-битных отрезков шифртекста с помощью подстановок, реализуемых с помощью S-блоков (S – substitution). К. Шеннон доказал [7], что эффективное преобразование блоков шифртекста может быть достигнуто путем чередования линейных операций перестановок битов, сохраняющих число значащих битов, и нелинейных операций, размножающих значащие биты, и реализуемых S-блоками<sup>2</sup>. Для линейных операций в DES используются P-блоки (P – permutation), а в ГОСТе – циклический сдвиг [2]. Из сказанного следует, что между двумя криптоалгоритмами много общего, что неудивительно, если учесть, что ГОСТ являлся российским ответом на DES [8, 9].

Целью криптопреобразований, примененных в обоих алгоритмах, является достижение лавинного эффекта, заключающегося в том, что каждый бит шифртекста зависит от каждого бита открытого текста и каждого бита ключа. В DES для этого нужно 5 циклов, тогда как в ГОСТ – 8. Однако ГОСТ состоит из 32 циклов, а DES – только из 16. После 8-ми циклов в DES наблюдается пик лавинного эффекта – каждый бит шифртекста является случайной функцией всех битов открытого текста и ключа. Успешные атаки на DES с тремя, четырьмя и шестью циклами подтвердили значение лавинного эффекта [2].

<sup>2</sup> В книге Н. Дж. А. Слоана «Коды, исправляющие ошибки, и криптография» (М.: Мир, 1983) отмечается, что подобное чередование операций эквивалентно «преобразованию пекаря», широко известному в нелинейной динамике.



Выше уже приводились данные результатов атак на блочные криптоалгоритмы, в том числе и на алгоритм RC5, отличающийся от DES длиной блока (от 32 до 128 бит), ключа (от 0 до 2048 бит) и числом циклов преобразований (от 0 до 255), а также другими деталями [2]. По этим же данным можно сделать заключение о значительном влиянии на криптостойкость длины ключа. Продолжительность атаки на RC5 с длиной блока 32 бита и числом циклов 12 варьировалась от 313 ч (длина ключа 48 бит) до 3.5 ч (длина ключа 40 бит).

Итак, криптостойкость алгоритма шифрования определяется в основном двумя элементами: ключом и таблицами подстановок [2], исследованию которых и будет посвящено дальнейшее изложение.

## 2. Блоки подстановок

Блоки подстановок или S-блоки представляют собой таблицы десятичных чисел от 0 до 15 размером 1x16 (ГОСТ) или 4x16 (DES). Это связано с особенностями преобразований. В ГОСТ на вход S-блоков поступает 32-битный отрезок шифртекста, который разбивается на восемь 4-битных отрезков по числу S-блоков, которые затем каждый по отдельности преобразуются на своем блоке и на выходе снова собираются в 32-битный отрезок. На вход S-блоков DES поступает не 32, а 48-битный отрезок шифртекста, полученный из 32-битного после перестановки с расширением [2]. Он разбивается на восемь 6-битных отрезков, каждый из которых поступает на вход одного из восьми S-блоков, представляющих аналогичные таблицы размером 4x16. Два крайних бита 0-й и 5-й, записанные в десятичной системе, определяют номер строки таблицы, по которой преобразуются биты с 1-го по 4-й. Объединенные преобразованные 32-битные отрезки поступают на вход P-блоков.

Отличие S-блоков DES и ГОСТ еще и в том, что для DES S-блоки являются частью стандарта, т.е. не меняются от сеанса к сеансу, тогда как стандарт ГОСТ не определяет способ генерации S-блоков. Однако общим для обоих криптоалгоритмов является то, что способ составления S-блоков является их секретной частью.

Очевидно, что от качества S-блоков зависит качество всего криптоалгоритма и, в первую очередь, его криптостойкость.

Приведем для дальнейшего исследования S-блоки ГОСТ [1, 2] (табл. 1), используемые в приложениях ЦБ РФ [1].

Таблица 1

S1	4	10	9	2	13	8	0	14	6	11	1	12	7	15	5	3
S2	14	11	4	12	6	13	15	10	2	3	8	1	0	7	5	9
S3	5	8	1	13	10	3	4	2	14	15	12	7	6	0	9	11
S4	7	13	10	1	0	8	9	15	14	4	6	12	11	2	5	3
S5	6	12	7	1	5	15	13	8	4	10	9	14	0	3	11	2
S6	4	11	10	0	7	2	1	13	3	6	8	5	9	12	15	14
S7	13	11	4	1	3	15	5	9	0	10	14	7	6	8	2	12
S8	1	5	13	0	5	7	10	4	9	2	3	14	6	11	8	12

Целью настоящей публикации является попытка выяснить, почему S-блоки, реализуемые криптоалгоритмами DES и ГОСТ, имеют именно такой вид, а не какой-либо другой, иными словами, попытаться хотя бы в общих чертах охарактеризовать алгоритм их выбора из множества всех возможных подстановок из 16-ти элементов.

Долгое время эта тема была предметом многочисленных обсуждений, и в 1992 г. IBM раскрыла секрет конструирования S-блоков. Однако указанные попытки не прекращались даже после этого, поскольку рекомендации IBM носили эвристический характер [1]. В [1] рассматриваются различные предложения по выработке S-блоков – от «ручного» конструирования на основе интуиции до случайного выбора или выбора на основе строгой математической теории. Анализ S-блоков, основанный на предположении об их случайном выборе, таит опасность, связанную с тем, что «проблема усложняется из-за способности человеческого сознания находить в случайных данных структуры, которые в действительности не являются структурами» [1].

В любом случае создатели криптоалгоритмов исходили из условия их устойчивости против известных методов криптоанализа и возможностей вычислительной техники [1].

В литературе [1, 2] встречаются утверждения, что выбор S-блоков ГОСТ осуществляется случайным образом. Как показано ниже, есть определенные свидетельства против этого.



Другой не менее важной целью является оценка качества S-блоков. Общеизвестно, что существуют «слабые» S-блоки. Тривиальный пример – тождественная подстановка, не меняющая 4-битный отрезок шифртекста. Другой не столь очевидный пример приведен в статье [10]:

$$S = (9, 8, 3, 10, 12, 13, 7, 14, 0, 1, 11, 2, 4, 5, 15, 6). \quad (1)$$

Его слабость становится явной, если записать его в двоичном виде в форме таблицы (табл.2).

Таблица 2

i	0000	0001	0010	0011	0100	0101	0110	0111
S(i)	1001	1000	0011	1010	1100	1101	0111	1110
i	1000	1001	1010	1011	1100	1101	1110	1111
S(i)	0000	0001	1011	0010	0100	0101	1111	0110

Здесь  $i$  – двоичное значение преобразуемого элемента,  $S(i)$  – его значение после подстановки. Из табл. 2 видно, что преобразование оставляет два бита из четырех неизменными<sup>3</sup>. Кроме этого таблица замен может содержать обходные пути других типов, позволяющие расшифровать сообщение более эффективным образом, чем полным перебором по возможным значениям ключа. Автор [10] высказывает пессимистическое утверждение о том, что не существует способа отсеять слабые таблицы подстановок.

### 3. Некоторые физические аналогии. Явление возврата

В дальнейшем мы будем рассматривать только S-блоки сами по себе безо всякой связи с криптоалгоритмами. Для этого есть все основания, поскольку подстановки – один из объектов дискретной математики и всегда представляли интерес для исследователей. Спектр приложений этих исследований весьма обширен [11].

Подстановка на множестве  $N$  элементов – это отображение данного множества на себя, и нас будут интересовать свойства этого

отображения. Для того чтобы не погрязнуть в абстракциях, мы будем, по возможности, прибегать к геометрическим иллюстрациям. Первое понятие из теории отображений, которое мы проиллюстрируем на примере S-блоков, это понятие возврата. А. Пуанкаре доказал теорему о возврате для непрерывных отображений [12], из которой следует, что с течением времени траектория динамической системы, совершающей движение в ограниченной области фазового пространства, вернется в сколь угодно малую окрестность начальной точки. Для дискретных систем это утверждение становится очевидным, поскольку точки фазового пространства характеризуются конечным объемом  $V$  и число шагов (аналог времени) до возврата оценивается сверху как отношение объема фазового пространства системы (многомерного в общем случае) к  $V$ . Проиллюстрируем это явление на примере отображения последовательности натуральных чисел от 0 до 15 с помощью слабого S-блока (1) (рис. 1). Программа выполняет итерации отображения (1), т.е.  $y_i = S^i(x)$ ,  $x = \{0, 1, \dots, 15\}$ ,  $i$  – номер итерации.

Из рис. 1 видно, что через 4 шага итераций отображаемая последовательность пришла в исходное состояние, т.е. произошло явление возврата. Легко понять, что это связано с тем, что максимальная длина циклов, на которые может быть разложена подстановка (1), равна 4, а длины остальных циклов являются делителями 4. Приведем полное разложение подстановки (1) на циклы:

$$S = C_4(9,8,1,0); C_4(3,10,11,2); C_4(7,14,15,6); C_2(12,4); C_2(13,5). \quad (2)$$

Здесь индекс означает длину цикла, а числа в скобках – элементы S-блока, преобразующиеся по данному циклу. Таким образом, мы показали, что слабость данного S-блока связана с небольшой длиной циклов, на которые он разлагается, точнее, с небольшой величиной их наименьшего общего кратного (НОК).

Представление подстановок с помощью циклов хорошо изучено [11]. Ниже мы еще воспользуемся этими результатами. Отметим, что в криптографической литературе им уделяется меньше внимания, чем они заслуживают.

<sup>3</sup> Само по себе число неизменных битов ни о чем не говорит. Например, для приведенной далее подстановки (2) число неизменных битов может достигать 3. Гораздо важнее, что это фиксированные биты (1-й и 2-й), а также то, что подстановка (1) не обладает свойством размешивания: группы 4, 5; 8, 9 и 12, 13 переходят в другие компактные группы.

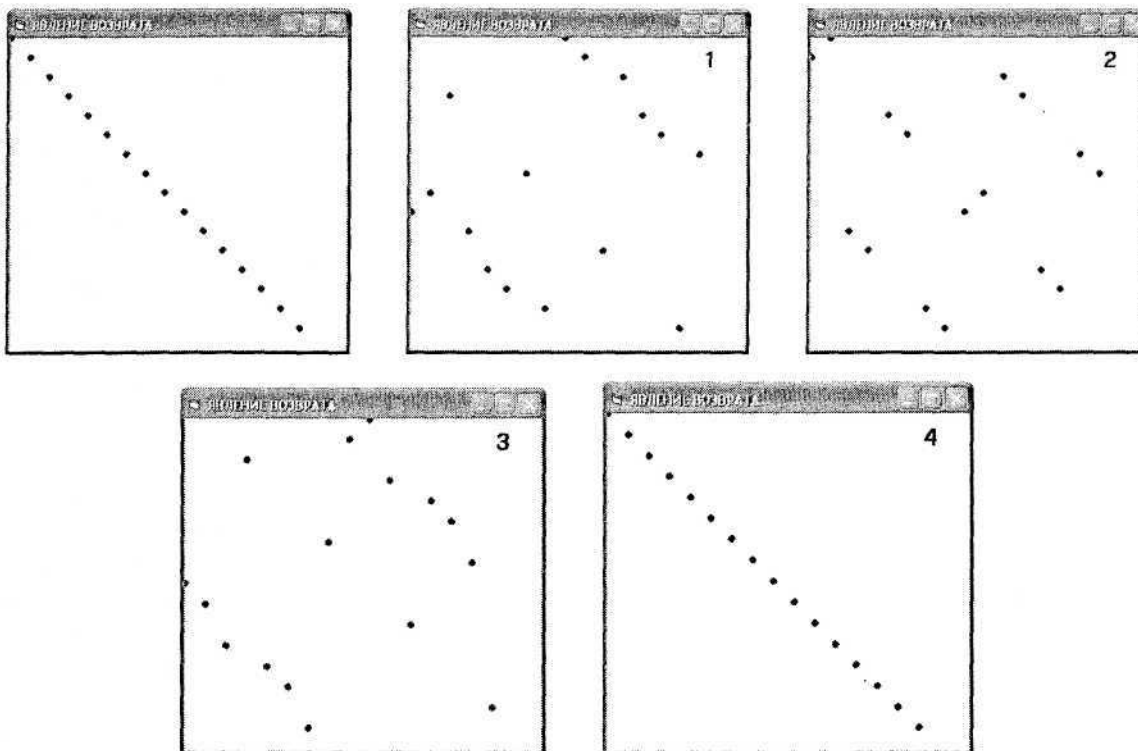


Рис. 1. Демонстрация явления возврата. Число в правом верхнем углу – номер итерации

#### 4. Некоторые физические аналогии. Эргодичность

Эргодичность – это свойство отображений, приводящее к равномерному заполнению траекторией всего фазового объема. В дальнейшем мы будем рассматривать отображения в фазовом пространстве, представляющем гиперкуб в восьмимерном пространстве с длиной ребра 15.

Существует еще одна характеристика отображений – перемешивание, которая оценивает степень разбегания траекторий динамической системы. В нелинейной динамике она оценивается с помощью так называемых показателей Ляпунова [12], характеризующих скорость разбегания двух изначально близких траекторий. Можно усмотреть аналогию между понятием перемешивания в динамике и понятием размешивающего преобразования в криптографии. Приведем цитату из [13]:

*«Под размешивающим преобразованием Шеннон понимает некоторое отображение векторного пространства на себя, при котором каждая или почти каждая его компактная область в отображении распределяется в большую, некокомпактную с точки зрения метрики, область...»*

*Рассматривая вектора как числа и применив подстановку:*

0 1 2 3 4 5 6 7  
7 3 1 5 4 2 3 0

*мы видим, что данное преобразование с большим основанием можно назвать размешивающим, т.к. компактная область 0 1 2 отображается в 7 3 1».*

Перемешивание – более сильное свойство динамической системы, чем эргодичность, в частности, система, обладающая перемешиванием, всегда эргодична. Системы с перемешиванием демонстрируют хаотическое, непредсказуемое поведение. В поисках аналогий можно было бы пойти еще дальше и сопоставить понятия динамического хаоса в динамике и лавинного эффекта в криптографии. Однако следует сказать, что прямой аналогии здесь нет, хотя бы потому, что отображения, реализуемые подстановками в криптографии и рассматриваемые здесь, всегда обратимы, тогда как хаотические системы в динамике демонстрируют необратимость.

В дискретных системах понятие ляпуновских показателей перестает работать. Если рассмотреть две изначально близкие тра-



ектории, принадлежащие одному циклу, то с ростом номера итерации расстояние между соответственными точками траекторий ведет себя случайным образом<sup>4</sup>. Поэтому вместо ляпуновских показателей можно характеризовать степень разбегания траекторий, а точнее, степень скорости, с которой стирается информация об их начальной близости с помощью автокорреляционной функции [14], вычисляя ее для зависимости расстояния от номера итерации (приложение).

Вернемся к эргодичности. Приведем ее определение [14]: «Если движение динамической системы эргодично, то относительное время, проведенное фазовой траекторией внутри любой области Г фазового пространства, равно относительному объему этой области и не зависит от выбора начальных условий. Иными словами, фазовая траектория эргодической системы будет равномерно и плотно заполнять всю область Г».

Ниже приведены результаты расчета траектории отображения, реализуемого итера-

циями S-блоков DAS в проекции на грани, соответствующие первым трем S-блокам (U = S11, V = S21, S = S31) (рис. 2).

На рис. 3, а, б показаны результаты расчета числа точек траектории, попавших в измерительный 8-мерный куб с центром, совпадающим с центром фазовой области в зависимости от L – размера половины ребра измерительного куба. Слева показаны объем измерительного куба (плавная кривая) и число точек траектории в нем (ломаная) в зависимости от L, а справа – их отношение. Начальная точка отображения имеет координаты 4, 15, 10, 7, 2, 12, 4, 2. Это соответствует длине траектории возврата C = 16016<sup>5</sup>.

Из полученных результатов следует, что точки траектории распределены не равномерно, а слоями, параллельными граням кубического фазового объема. Слабая зависимость результата от числа итераций (за исключением числа точек отображения) позволяет говорить о равномерном заполнении точками всего фазового пространства, т.е. эргодичности S-блоков DES (по крайней мере, выбранных) (табл. 3).

Таблица 3

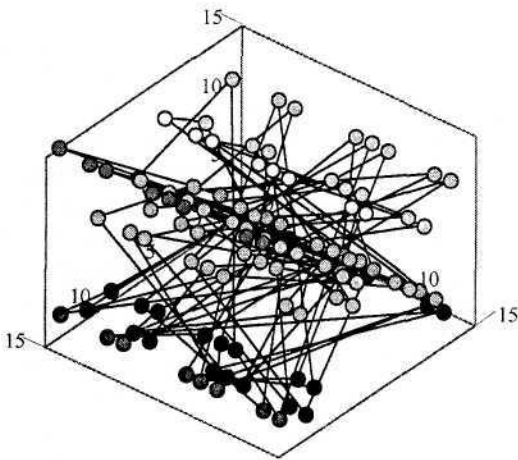
S-блоки DES

S11 =	14 4 13 1 2 15 11 8 3 10 6 12 5 9 0 7	S21 =	15 1 8 14 6 11 3 4 9 7 2 13 12 0 5 10	S31 =	10 0 9 14 6 3 15 5 1 13 12 7 11 4 2 8	S41 =	7 13 14 3 0 6 9 10 1 1 2 8 5 11 12 4 15	S51 =	2 12 4 1 7 10 11 6 8 5 3 15 13 0 14 9	S61 =	12 1 10 15 9 2 6 8 0 8 13 3 4 14 7 5 11	S71 =	4 11 2 14 15 0 8 13 3 9 7 5 10 6 1	S81 =	13 2 8 4 6 15 11 1 10 9 3 14 5 0 12 7
-------	--	-------	--	-------	--	-------	---	-------	--	-------	---	-------	--	-------	--

Примечание. Использованы первые строки каждого S-блока.

<sup>4</sup> Если соответствующий S-блок достаточно велик, например имеет число входов-выходов равнос 256 (см. приложение), и обладает хорошим рассеиванием, то этим свойством можно воспользоваться для генерации псевдослучайных последовательностей.

<sup>5</sup> Чтобы сравнивать число точек и объем гиперкуба, каждой точке отображения приписывается «объем»  $15^8 / (C + 1)$ . На рисунках видны следствия этой довольно произвольной процедуры – ломаная кривая пересекает плавную кривую и может сложиться впечатление, что суммарный «объем» точек в гиперкубе превосходит объем всего гиперкуба.



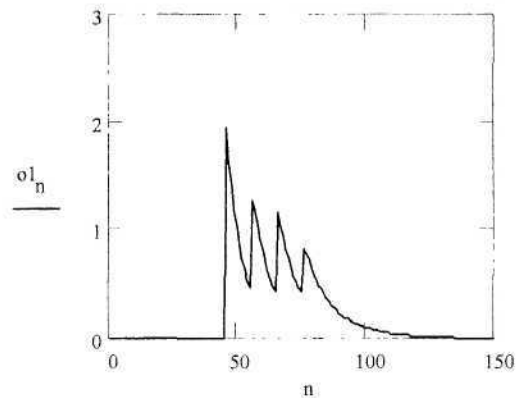
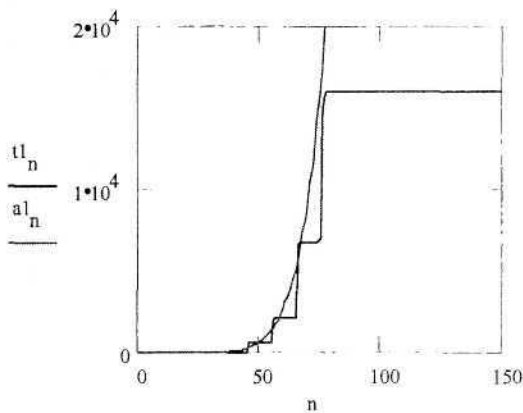
U, V, S

Рис. 2. Трехмерная проекция фазовой траектории для элементов S-блоков с номерами (1 0 0 0 0 0 1). Длина траектории возврата равна 16016

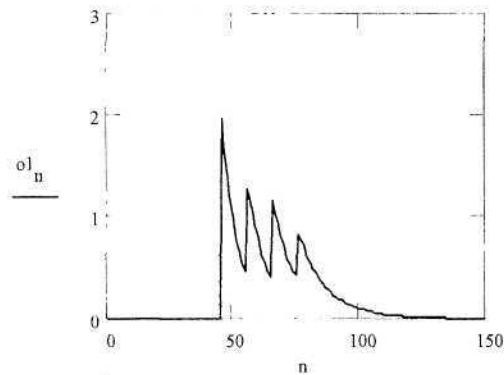
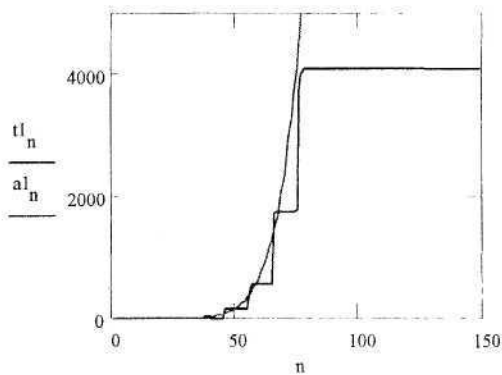
На рис. 4, а, б показаны результаты расчета для тех же S-блоков, но для другой начальной точки с координатами 14, 15, 10, 7, 2, 12, 4, 13, которой соответствует траектория возврата с длиной  $C = 1232$ .

Сравнение рис. 3, а и 4, а позволяет утверждать, что сказанное выше о равномерности заполнения точками отображения фазового пространства не зависит от конкретной траектории отображения и, в частности, от длины траектории возврата<sup>6</sup>.

Другая особенность отображения – отсутствие точек в центре куба, связана с многомерным характером траектории и согласуется с известной особенностью многомерных объектов – вклад в их полный объем дают в основном периферические области.



а



б

Рис. 3. Результаты исследования эргодичности для S-блоков DES. Длина ребра измерительного куба равна  $2L$ ,  $L = n/10$ ,  $0 < n < 150$ . Точки перестают накапливаться при  $n > 75$ , т.е. после прохождения границы фазового объема. Число итераций  $C = 16016$  (а) и  $C = 4096$  (б)

<sup>6</sup> Необходимо помнить, что каждая данная совокупность S-блоков генерирует разные траектории. Исключение составляет случай, когда все подстановки в S-блоках полноцикловые.

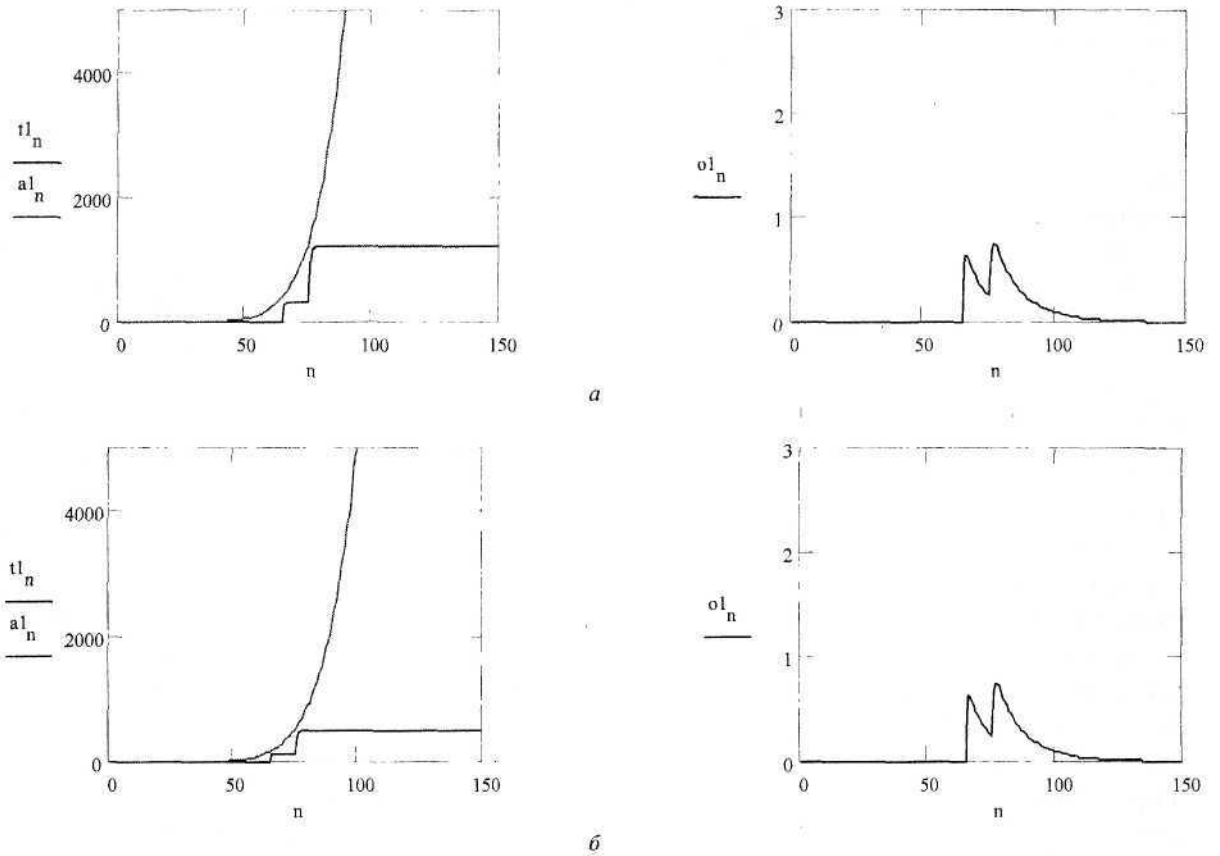


Рис. 4. Расчет для тех же S-блоков DES с начальной точкой, соответствующей номерам элементов (0 0 0 0 0 0 0). Число итераций  $C = 1232$  (а) и  $C = 512$  (б)

Для сравнения на рис. 5 приведем результаты аналогичного расчета для одного S-блока – S31, реализующего полноцикло-

вую подстановку с длиной траектории возврата 16.

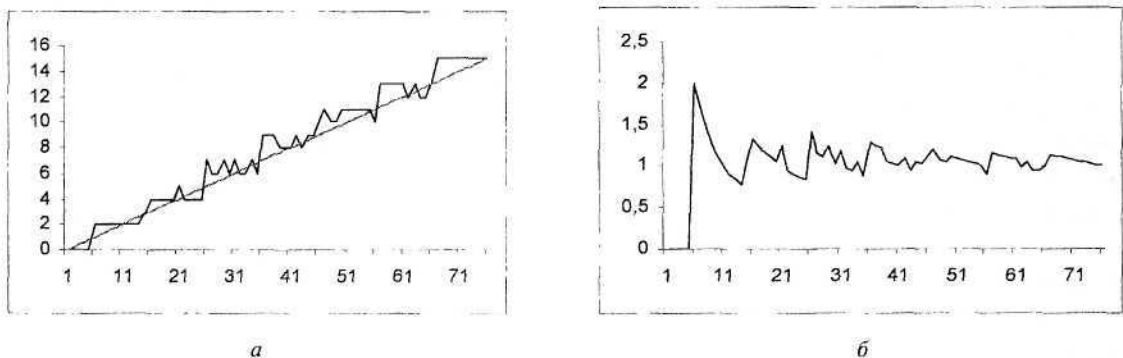


Рис. 5. Результаты исследования эргодичности для блока S31: а –  $2L$  – длина ребра измерительного куба (отрезка) – прямая;  $N$  – число точек в кубе – ломаная; б –  $N/2L$  – их отношение. По оси абсцисс отложена величина  $L/10$

### 5. Качество подстановок

Вопрос о качестве подстановок, реализуемых S-блоками блочных шифров, является одним из основных. Хороший S-блок (наряду с хорошим ключом) должен обеспечи-

вать выполнение требований, предъявляемых к криптостойкости шифра. Очевидно, в шифросистемы ГОСТ 28147-89 и DES заложены алгоритмы выработки S-блоков, удовлетворяющие этим требованиям. Однако они



не раскрываются. Вместе с тем нет гарантии, что в некоторых реализациях упомянутых криптосистем не используются ослабленные S-блоки, наподобие (1). Поэтому исследователи-криптографы пытаются найти правила, позволяющие а priori оценивать качество подстановок, используемых в S-блоках.

В [9] приводятся некоторые, достаточно сложные, рассуждения, целью которых является нахождение лучших возможных подстановок. Для 16-ти элементов приведен явный вид подстановки, названной логарифмической:

14 12 3 7 9 15 8 13 0 6 2 10 5 4 1 1 (3)

Приведем цитату из [9]:

*«Это был, пожалуй, мой самый красивый математический результат. Но, к большому сожалению, логарифмические подстановки так и не нашли достойного применения в криптографии. Почему? Да очень просто – их мало. Помните фразу про долговременные ключи-подстановки в дисковых шифраторах: “Их не опробуют. Их покупают”. Если в схемы типа “Ангстрем-3” мы будем ставить только логарифмические подстановки, то опробование всевозможных вариантов подобных подстановок сведется к опробованию всего лишь трех элементов:  $q$  – примитивного элемента в поле Галуа  $GF(257)$ ,  $r$  – произвольного ненулевого элемента поля  $GF(257)$  и  $r$  – произвольного элемента из  $Z/256$ . Это – копейки, совершенно ничтожная, по криптографическим меркам, величина. Если же выбирать подстановку случайно и равновероятно из всей симметрической группы  $S_{256}$  (т.е. группы подстановок на множестве из 256 элементов. – Ю.З.), то общее число опробуемых вариантов будет совершенно астрономической величиной  $256!$ , намного превосходящей психологически недостижимую в криптографии величину  $10^{100}$ .*

*Но для шифров на новой элементной базе логарифмические подстановки позволили полнее представить общую картину того “лавинного эффекта”, к достижению которого так стремятся криптографы всего мира».*

Эта цитата позволяет хотя бы приблизительно оценить сложность поставленной задачи и получить представление о методах,

которые применяют специалисты-криптографы для ее решения.

Выше приводился пример слабого S-блока (1) и было высказано соображение, что его слабость связана с наличием коротких циклов, на которые разлагается реализуемая им подстановка. Можно было бы предположить, что полноцикловые подстановки, т.е. подстановки, состоящие из одного цикла максимальной длины, заведомо не будут слабыми и будут приводить к преобразованиям с хорошими размешивающими свойствами. То, что это не так, легко видно из простого примера полноциклового подстановки, не удовлетворяющей требованию размешивания:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 0

Другой пример связан с упоминавшимися выше логарифмическими подстановками. Среди них могут быть полноцикловые, как, например, (3), и состоящие из более чем одного цикла (М. Масленников, частное сообщение).

Логарифмические подстановки, как уже было сказано, не находят широкого применения из-за того, что их мало и, следовательно, они могут быть легко вычислены.

Полноцикловые подстановки даже с хорошим размешиванием опасно применять еще и по следующим соображениям. Рассмотрим набор из восьми S-блоков, реализующих исключительно полноцикловые подстановки. Поскольку они оперируют независимо друг от друга, естественно перенести рассмотрение в восьмимерное пространство. Поставим вопрос о длине траектории возврата для данного варианта. Ответ очевиден – 16. Т.е. через 16, 32 и т.д. итераций преобразуемый текст восстановится. Конечно, этот результат получен в пренебрежении всеми остальными деталями криптоалгоритма, но это и позволило подчеркнуть слабую сторону полноцикловых подстановок.

Теперь посмотрим, какие подстановки применяют на практике (на примере ГОСТа). Для этого разложим S-блоки ГОСТ на циклы:

$S1 \rightarrow C_2(10, 1); C_{14}(0, 4, 13, 15, 3, 2, 9, 11, 12, 7, 14, 5, 8, 6)$

$S2 \rightarrow C_2(1, 11); C_{14}(0, 14, 5, 13, 7, 10, 8, 2, 4, 6, 15, 9, 3, 12)$





- $S3 \rightarrow C_4(0, 5, 3, 13); C_4(4, 10, 12, 6);$   
 $C_8(8, 14, 9, 15, 11, 7, 2, 1)$   
 $S4 \rightarrow C_2(11, 12); C_3(8, 14, 5);$   
 $C_{11}(7, 15, 3, 1, 13, 2, 10, 6, 9, 4, 0)$  (4)  
 $S5 \rightarrow C_2(11, 14); C_2(9, 10); C_6(5, 15, 2,$   
 $7, 8, 4); C_6(6, 13, 3, 1, 12, 0)$   
 $S6 \rightarrow C_2(14, 15); C_{14}(4, 7, 13, 12, 9, 6, 1,$   
 $11, 5, 2, 10, 8, 3, 0)$   
 $S7 \rightarrow C_3(13, 8, 0); C_4(15, 12, 6, 5);$   
 $C_9(11, 7, 9, 10, 14, 2, 4, 3, 1)$   
 $S8 \rightarrow C_3(5, 7, 4); C_6(13, 11, 14, 8, 9, 2);$   
 $C_7(1, 15, 12, 6, 10, 3, 0)$

Видно, что создатели алгоритма использовали самые разные подстановки. Причина этого уже понятна. Отметим, что, в отличие от DES, среди циклов ГОСТ нет единичных,

что, конечно, усиливает криптоалгоритм. Попробуем оценить длину траектории возврата. Для этого найдем НОК длин всех входящих в разложение (4) циклов. Полученное число равно 5544 – это верхняя граница искомой величины. Представляется разумным предположить, что разработчики шифра стремились выбирать подстановки так, чтобы сделать это число максимально большим. Однако вопрос, каким образом они этого достигали, остается открытым.

Наконец, попытаемся представить качество подстановок визуально с помощью отображения первого возвращения, демонстрирующего зависимость  $S(S(i))$  от  $S(i)$  ( $i$  – номер элемента) для двух представленных выше подстановок и реального S-блока (рис. 6).

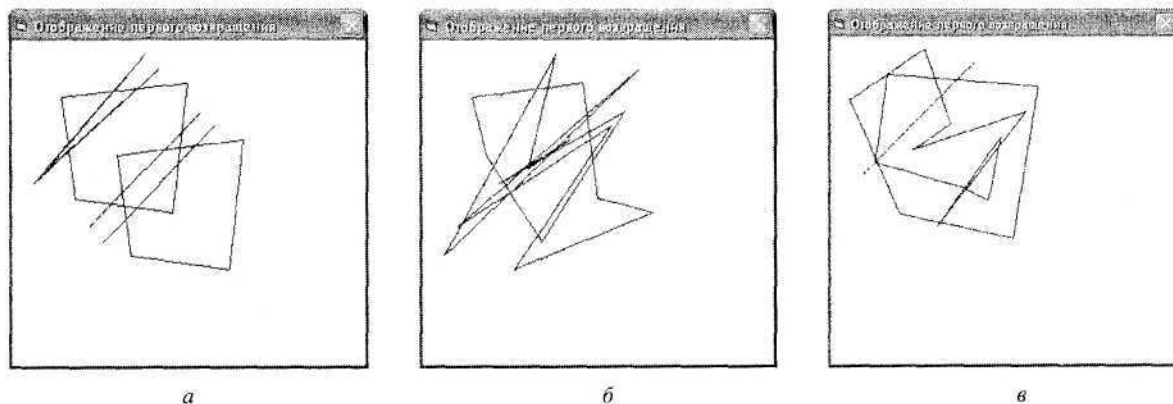


Рис. 6. Отображение первого возвращения:  $a$  – для подстановки (1);  $b$  – для (2);  $v$  – для S1-блока ГОСТ

Отличие отображений бросается в глаза. Отображение  $a$  резко отличается от отображений  $b$  и  $v$  своей регулярностью, связанной с плохими криптографическими свойствами подстановки (1). С другой стороны, на каждом отображении четко просматривается цикловая структура соответствующей подстановки. Особенно четко она просматривается на рис. 6,  $a$ . Две прямые соответствуют двум циклам длины 2, три четырехугольника – трем циклам длины 4. На основании рис. 6 можно попытаться сформулировать геометрический критерий отбора «хороших» S-блоков.

## 6. Статистика S-блоков

Представление подстановок с помощью циклов позволяет исследовать их статистические характеристики и сопоставить с ха-

рактеристиками реальных S-блоков. Результаты статистического исследования подстановок мы заимствуем у Д. Кнута [11].

Если  $A$  – число циклов в подстановке<sup>7</sup> из  $n$  элементов, то его средние характеристики суть:

$$\begin{aligned} \min A &= 1, \quad \text{ave } A = H_n, \\ \max A &= n, \quad \text{dev } A = (H_n - H_n^2)^{1/2}. \end{aligned} \quad (5)$$

Здесь  $\text{ave}$  и  $\text{dev}$  – среднее значение и среднеквадратичное отклонение соответственно (в обозначениях Д. Кнута) и

$$H_n^m = \sum_{1 \leq k \leq n} \frac{1}{k^m}, \quad H_n = H_n^1. \quad (6)$$

Кроме того, подстановка из  $n$  предметов имеет  $k$  циклов с вероятностью, равной

<sup>7</sup> Д. Кнут называет их перестановками.



$|S(n, k)| / n!$ , где  $(-1)^{n-k} S(n, k)$  – числа Стирлинга первого рода.

Можно также задать вопрос о средней длине одного цикла. Согласно (5), общее число циклов во всех  $n!$  перестановках есть  $n!H_n$  (поскольку оно равно среднему числу циклов, взятому  $n!$  раз). Отсюда можно получить выражение для средней длины произвольного цикла в виде  $n/H_n$ .

Отсюда же следует выражение для вероятности появления цикла длины  $k$  в подстановке из  $n$  элементов. Поскольку общее число циклов длины  $k$  в подстановке из  $n$  элементов равно  $n!/k$ , а общее число циклов

есть  $n!H_n$ , то искомая вероятность равна их отношению, т.е.  $1/kH_n$ .

Подсчитаем вероятность появления полноцикловых подстановок на множестве из 16 элементов. Она равна  $1/16H_{16} = 0.0185$ .

На рис. 7 представлены статистические результаты по длинам циклов. На их основе можно решить вопрос о характере отбора S-блоков из всего множества подстановок. Уже беглого взгляда достаточно, чтобы понять, что для ГОСТ отбор делался не случайно и не равномерно. Чтобы получить точный ответ, приведем результаты корреляционного анализа представленных данных.

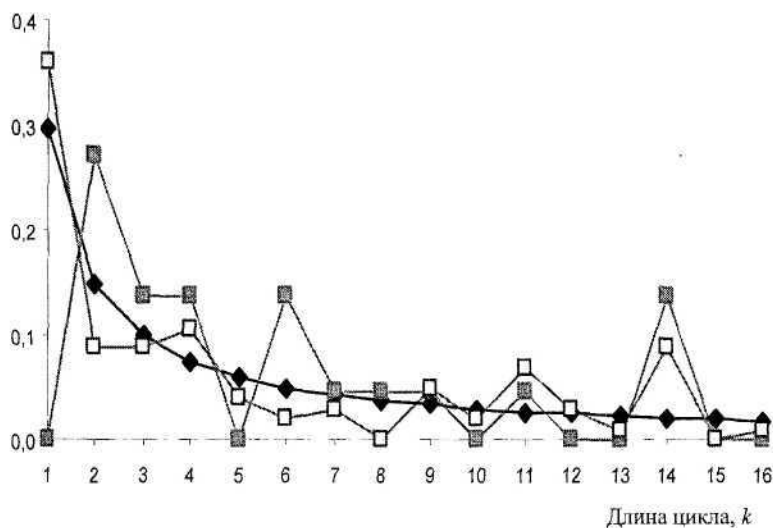


Рис. 7. Статистические характеристики длин циклов: —◆— вероятность  $P(k)$ ; —■— частота ГОСТ; —□— частота DES

Формула для оценки коэффициента корреляции двух случайных величин  $x$  и  $y$  имеет вид [15]:

$$r_{xy} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\left[ \sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2 \right]^{1/2}} = \frac{\sum_{i=1}^N x_i y_i - N\bar{x}\bar{y}}{\left[ \left( \sum_{i=1}^N x_i^2 - N\bar{x}^2 \right) \left( \sum_{i=1}^N y_i^2 - N\bar{y}^2 \right) \right]^{1/2}} \quad (7)$$

Здесь  $x_i$  и  $y_i$  – вероятность появления в подстановке цикла длиной  $i$  и частота этого события для конкретного криптоалгоритма

(DES или ГОСТ),  $\bar{x}$  и  $\bar{y}$  – их среднее и выборочное среднее значения,  $N = 16$  – число испытаний (т.е. различных циклов) в выборке.

Область принятия гипотезы о нулевой корреляции (т.е. отсутствии корреляции) имеет вид [15]:

$$-z_{\alpha/2} \leq \sqrt{N-3} \cdot w \leq z_{\alpha/2}, \quad w = \frac{1}{2} \ln \frac{1+r_{xy}}{1-r_{xy}} \quad (8)$$

где  $z$  – стандартная, нормально распределенная величина. Если значение окажется вне этого интервала, то это будет признаком наличия статистической корреляции с уровнем значимости  $\alpha$ .



Выполняя расчеты, получаем следующие значения. Для ГОСТ:  $r_{xy} = 0.203783$  и  $w = 0.206676$ , а для DES –  $r_{xy} = 0.916768$  и  $w = 1.568383$ . Сразу бросается в глаза, что значение коэффициента корреляции для DES более чем в четыре раза превышает аналогичное значение для ГОСТ. Задавая уровень значимости  $\alpha = 5\%$  и беря значение  $z_{\alpha/2}$  из таблиц [15, с.500] равным 1.96, получим, что гипотеза о нулевой корреляции с указанным уровнем значимости должна быть отвергнута

для DES, поскольку  $w(N-3)^{1/2}$ , равное 5.6549, не попадает в интервал  $\pm 1.96$ . В то же время для ГОСТ эта гипотеза с тем же уровнем значимости должна быть принята, так как значение  $w(N-3)^{1/2}$ , равное 0.7452, оказывается внутри этого интервала.

Проведем аналогичные расчеты для распределения S-блоков по числу циклов, на которые они разлагаются, исходя из публикуемых данных (рис.8).

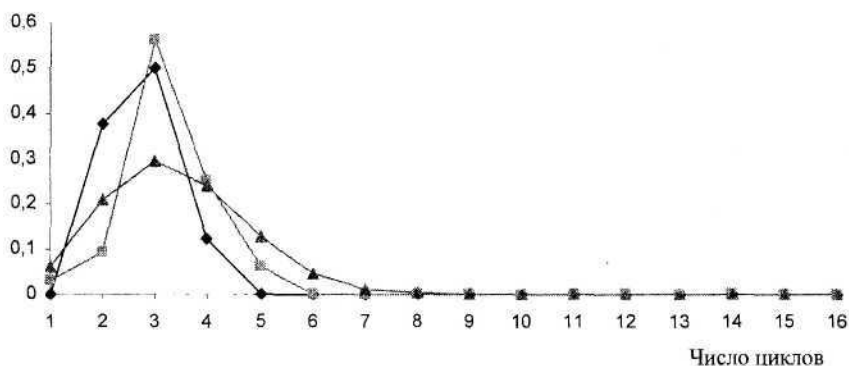


Рис. 8. Статистические данные по числу циклов в S-блоках: —◆— ГОСТ; —■— DES; —▲— вероятность

Выполняя расчеты по формулам (7) и (8), получим: для ГОСТ – коэффициент корреляции  $r_{xy} = 0.852265$ ,  $w = 1.264372$  и  $w(N-3)^{1/2} = 4.558759$ , а для DES –  $r_{xy} = 0.804818$ ,  $w = 1.112141$  и  $w(N-3)^{1/2} = 4.009883$ . Задавая тот же уровень значимости 5%, получим, что гипотеза о нулевой корреляции как для ГОСТ, так и для DES должна быть отвергнута, поскольку полученные значения не попадают в интервал  $\pm 1.96$ . Говоря другими словами, если за фактор оценки выбора S-блоков взять число циклов в них, то полученный результат говорит в пользу случайного и равновероятного отбора S-блоков из множества всех подстановок.

Следует сказать, что этот результат, как и приведенный выше, не вносит окончательной ясности в вопрос о случайности (или неслучайности) выбора S-блоков. В литературе встречаются утверждения обоего толка [1, 2]. Ввиду того что упомянутые утверждения представляют собой мнения, не подкрепленные ссылками на иные источники, кроме

других мнений [16] (Бёрд Киви, частное сообщение), следует, по-видимому, относиться к ним с осторожностью, по крайней мере, в отношении публикуемых S-блоков.

Можно привести некоторые дополнительные соображения о неслучайности выбора S-блоков, привлекая для этой цели закон больших чисел [15]. Согласно ему разности выборочных средних случайной величины, взятые по некоторым выборкам мощности  $N_1$  и  $N_2$ , и математического ожидания этой величины относятся как  $(N_1/N_2)^{-1/2}$ . Чтобы применить закон больших чисел к данной ситуации, надо пренебречь отличиями алгоритмов выбора S-блоков DES и ГОСТ и считать их наборы двумя случайными выборками разной мощности  $N_2 = 32$  и  $N_1 = 8$ . Тогда вышеупомянутое отношение должно быть равно 2. Однако в действительности, как легко проверить с помощью вышеприведенных формул, оно близко к 4, причем это справедливо как для числа циклов в перестановках, так и для их длин.



Следует сказать, что высказывание А. Чморя [2] о случайном выборе S-блоков ГОСТ заимствовано им у Б. Шнайера [1], что снимает с него ответственность за необоснованные «домыслы» (Бёрд Киви, частное сообщение). Что же касается первоисточника [1], то, скорее всего, это утверждение – результат слишком вольного перевода.

#### Заключение

В статье представлены некоторые результаты исследования важных элементов блочных криптографических систем – блоков подстановок, так называемых S-блоков, методами, которые заимствованы из нелинейной динамики сложных систем. Эти методы используют понятия точечного отображения, эргодичности, перемешивания и др. Сами эти понятия не являются чем-то чуждым в криптографии, например, понятие отображения с перемешиванием близко к понятию размешивающего преобразования. Новым является использование визуального представления преобразований, осуществляемых S-блоками, что позволяет наглядно представить процесс оценки качества и отбора «хороших» с криптографической точки зрения подстановок.

Другим результатом статьи является статистический анализ S-блоков, выполненный на основе разложения их на циклы. Это позволяет более обоснованно, чем это делалось ранее, решить вопрос о выборе S-блоков из всего множества подстановок.

Возможно, в дальнейшем этот подход в сочетании с традиционными для криптографии методами позволит по-новому взглянуть на проблемы блочных криптографических систем и получить новые полезные результаты.

Автор выражает благодарность М. Масленникову за обсуждения статьи, а также студентам ФГОУ ВПО ПАГС им. П.А. Столыпина и филиала в г. Балакове, обучающимся по специальности «Прикладная информатика», О. Елистратовой и М. Конинской за активное и полезное участие в работе.

#### Список литературы

1. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы и исходные тексты на языке C / Пер. с англ. М.: Триумф, 2002. 816 с.
2. Чморя А.Л. Современная прикладная криптография. М.: Гелиос АРВ, 2002.

3. Указ Президента РФ № 334 от 05.04.95 «О мерах по соблюдению законности в области разработки, производства, реализации и эксплуатации шифровальных средств, а также предоставление услуг в области шифрования информации» // Собрание законодательства РФ. 1995. №29.
4. Federal Information Processing Standards Publication (FIPS PUB) 46-1, Data Encryption Standard, Reaffirmed 1988 Jan 22 (supersedes FIPS PUB 46, 1977 Jan 15).
5. ГОСТ 28147-89 «Системы обработки информации. Защита криптографическая. Алгоритм криптографического преобразования». М.: Госстандарт СССР, 1989.
6. Бернет С., Пэйн С. Криптография. Официальное руководство RSA Security / Пер. с англ.; Под ред. А.И. Тихонова. М.: Бином, 2002.
7. Шеннон К. Работы по теории связи и кибернетике. М.: Инostr. лит., 1963. 830 с.
8. Масленников М. Практическая криптография. СПб.: БХВ-Петербург, 2003.
9. Масленников М. Криптография и свобода // <http://mikhailmasl.livejournal.com/4852.html>
10. Винокуров А.Ю. Как устроен блочный шифр // <http://algolist.manual.ru/defence/index.php>
11. Кнут Д. Искусство программирования: В 4 т. Т.1. Основные алгоритмы. М.: Мир, 1976.
12. Заславский Г.М. Стохастичность динамических систем. М.: Наука, 1984.
13. Щербаков А.Ю., Домашев А.В. Прикладная криптография. Использование и синтез криптографических интерфейсов. М.: Русская редакция, 2003.
14. Лоскутов А.Ю., Михайлов А.С. Введение в синергетику. М.: Наука, 1990.
15. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. М.: Мир, 1989.
16. [http://en.wikipedia.org/wiki/Data\\_Encryption\\_Standard](http://en.wikipedia.org/wiki/Data_Encryption_Standard) (англоязычная версия), <http://ru.wikipedia.org/wiki/DES> (русскоязычная версия).

#### Приложение

##### Тестирование S-блоков хеш-функций методами нелинейной динамики

Хеш-функции применяются в криптографии для целей шифрования паролей, создания цифровой подписи и т.д. [17]. Среди направлений разработки стойких хеш-функций можно выделить использование симметричных блочных алгоритмов [18], одним из основных элементов алгоритма которых является преобразование подстановки, выполняемое с помощью S-блоков, подобных рассмотренным в настоящей статье. Применим изложенную выше методику для оценки качества используемых в реальных алгоритмах хеширования S-блоков, исследуя эргодичность реализуемых ими преобразований. Для этого рассмотрим несколько заявок конкурса, проводимого Национальным институтом стандартов США (NIST), информация о котором была помещена на официальном сайте NIST [19] в декабре 2008 г.

Одна из заявок зарегистрирована под кодовым именем Abacus (заявитель – Neil Sholer). Не вдаваясь в детали, отметим, что в предлагаемом алгоритме используется S-блок 256\*256, преобразующий входные байты в выходные (256 – число входов и выходов). S-блок представлен в табл. П.1 (в шестнадцатеричной системе).



Таблица П.1

	00	01	02	03	04	05	06	07	08	09	0a	0b	0c	0d	0e	0f
00	e3	84	f0	d6	f9	f6	bc	90	85	7d	28	43	12	c0	e1	b4
10	55	e7	8c	87	42	e0	d9	27	78	ec	cb	07	aa	95	c1	3f
20	b2	dc	26	a7	1f	df	f3	54	d2	c7	24	3e	32	d1	56	c6
30	35	73	f7	7b	62	29	52	80	a9	ba	ab	e9	02	53	6a	e4
40	67	a0	8e	fb	9a	79	4e	8d	e5	4a	41	af	5a	5c	a6	6b
50	16	5e	e8	3c	9c	5b	88	76	15	f4	60	bd	83	98	8f	c8
60	09	68	0d	18	65	45	04	ce	7a	f2	39	c5	9e	f1	17	ef
70	38	21	94	86	69	37	f5	ed	36	66	cf	3b	63	4b	33	b6
80	ff	bc	11	5d	b3	2b	d3	d0	3a	96	77	7c	1c	c2	fe	0a
90	e3	25	4d	fc	89	de	30	23	64	81	d5	ae	70	db	e6	7e
a0	b0	6f	0f	d7	bf	9b	c4	74	b7	57	4f	58	10	2d	a4	b9
b0	a2	ad	61	eb	ac	1a	a3	d8	2c	5f	91	2f	72	31	b1	82
c0	49	da	0c	ca	00	a1	b5	75	6e	47	6d	13	19	93	20	05
d0	01	9f	1d	44	8a	1e	50	34	fa	9d	a8	8b	0b	4c	a5	2e
e0	71	f8	40	cd	99	fd	51	59	0e	2a	3d	92	14	48	6c	ea
f0	46	22	cc	06	d4	97	e2	1b	dd	7f	bb	c9	b8	03	ee	08

Старшие значащие биты расположены слева. S-блок читается слева направо и сверху вниз. Например, входу 02 соответствует выход f0, а входу 7e – выход 33. Ниже представлены результаты исследования эргодичности S-блока.

На рис. П.1 показана зависимость отношения числа точек в измерительном кубе  $n$  (отрезке) к длине отрезка  $2L$  в зависимости от  $L$ ,  $1 \leq L \leq 127$ . Центр куба расположен в точке 127. Кроме того, там же представлены результаты расчета автокорреляционной функции для расстояния двух изначально близких точек, принадлежащих одному циклу по формуле

$$Corr(j) = \frac{1}{C} \sum_{i=0}^{C-j} [r(i) - \langle r \rangle] \cdot [r(i+j) - \langle r \rangle], \quad (9)$$

где  $r$  – расстояние между точками,  $\langle r \rangle$  – среднее расстояние (по итерациям),  $i, j$  – номер итерации,  $C$  – длина цикла (траектории возврата). Начальные точки выбраны так, что  $r(0) = 1$ .

S-блок имеет три цикла (подстановка разлагается на три цикла) длины  $C = 12, 61$  и  $183$  ( $12 + 61 + 183 = 256$ ). Для каждого из них получается своя характеристика эргодичности. Для цикла длиной 12 график показывает насыщение величины  $n/2L$  порядка  $0.05 \approx 12/256 = 0.046875$ . Для цикла длиной 61 – насыщение порядка  $0.25 \approx 61/256 = 0.23828$ . Для цикла длиной 183 насыщение порядка  $0.7 \approx 183/256 = 0.71484$ .

Эти значения представляют собой не что иное, как значения инвариантной меры на соответствующих подмножествах, совпадающих с циклами подстановки. Это следует из определения понятия инвариантной меры [20], позволяющей выразить среднее значение произвольной функции  $g(x)$  на множестве элементов подстановки как среднее значение относительно инвариантной меры

$$\frac{1}{N} \sum_{i=0}^{N-1} g(x_i) = \frac{1}{N} \sum_j \sum_{i=0}^{N_j-1} g(f^i(x_{0j})) = \sum_j \rho_j \sum_{i=0}^{N_j-1} g(x_i). \quad (10)$$

Здесь  $N = 256$  – полное число элементов подстановки,  $N_j = 12, 61, 183$  ( $j = 1, 2, 3$ ) – число элементов каждого из циклов, на которые распадается подстановка  $f$ ;  $f^i$  –  $i$ -я итерация подстановки  $f$ ,  $x_{0j}$  – произвольный элемент подстановки, принадлежащий  $j$ -му подмножеству (циклу);  $\rho_j = N_j / N$  – значения инвариантной меры  $\rho$  на  $j$ -м подмножестве.

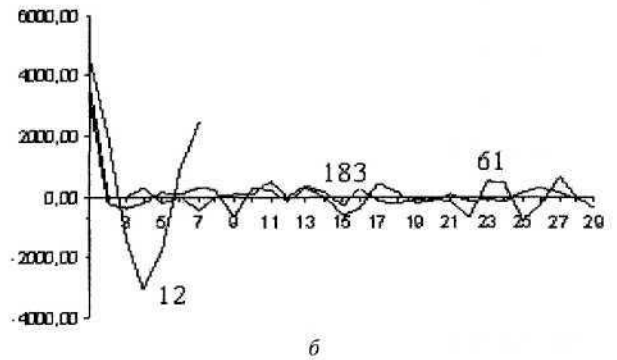
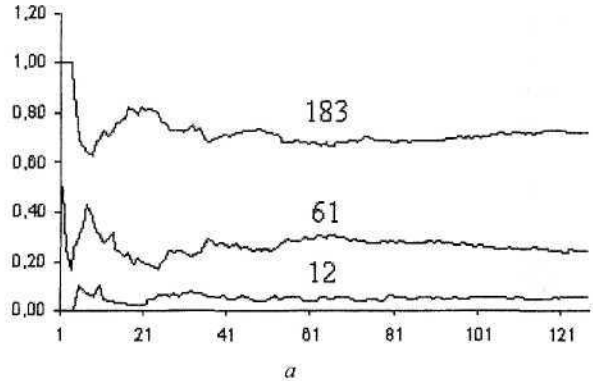


Рис. П.1. Результаты тестирования S-блока Abacus: а – результаты расчета эргодичности; б – автокорреляционная функция. Число у кривой означает длину соответствующей траектории возврата. По оси абсцисс отложена величина  $L/2$  (а) и номер итерации  $j$  (б)

Из анализа поведения автокорреляционной функции можно заключить, что переход от цикла длиной 183 к циклу длиной 61 приводит к значительному росту ее значений на итерациях с большим номером. Для цикла длиной 12 рост значений автокорреляционной функции имеет место при любых итерациях.

Рассмотрим характеристики другого кандидата – MCSSHA-3 (заявитель М. Масленников). Соответствующий S-блок представлен в табл. П.2

Таблица П.2

30	60	67	B5	43	EA	93	25	48	0D	18	6F	28	7A	FE	B6
D5	9C	23	86	52	42	F7	FD	F6	9B	EE	99	91	BC	2A	63
A1	A0	57	3C	39	D2	EC	71	45	CB	41	DC	0B	5B	C2	36
01	55	7D	FB	ED	83	8F	31	C0	4C	08	E3	9D	C1	D3	E9
B8	BD	AE	0F	E7	70	5A	EB	4D	29	F9	A9	3D	26	46	06
D0	50	A5	BE	66	90	F4	20	E4	33	27	E2	AB	EF	68	54
37	6A	DB	BB	D8	7B	69	C4	F2	BF	85	C7	A6	B4	9A	DD
72	34	E8	FC	D6	21	98	96	32	CA	49	B3	F3	97	8E	2F
00	B0	10	1A	77	38	CF	51	BA	1F	22	AC	62	89	76	C3
02	6E	2C	47	3A	5C	1B	56	8A	5D	03	16	74	58	79	09
D7	F5	0A	92	4F	87	CD	DA	8C	C9	9E	3B	12	6B	53	FF
80	B7	F8	D9	F1	5E	AF	E0	05	A4	14	2B	A3	CC	6C	7C
78	AA	95	84	61	A8	CE	13	88	FA	59	4E	B9	C8	4B	24
D1	07	94	2E	DF	B1	17	A2	1D	4A	C6	AD	15	19	35	7F
81	44	0C	9F	75	7E	D4	82	DE	F6	E1	2D	3E	73	11	8B
C5	A7	F0	6D	IC	64	0E	04	40	1E	8D	E5	3F	B2	65	5F

Результаты анализа показали, что подстановка имеет один цикл длиной 256. Это хорошо видно на рис. П.2.

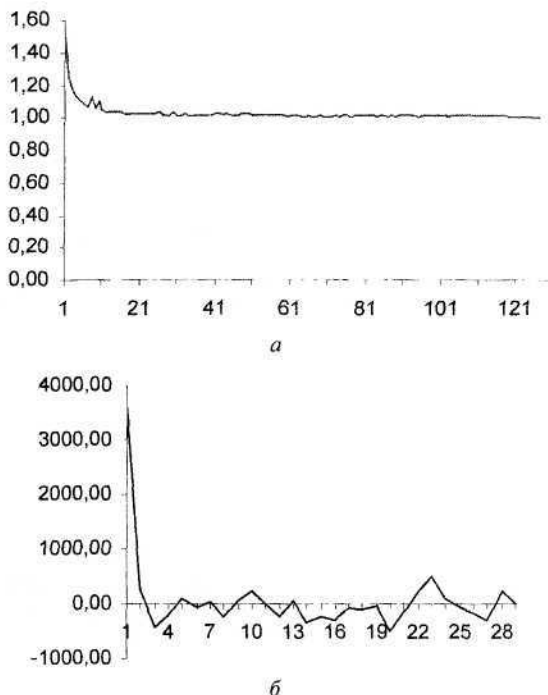


Рис. 11.2. Результаты тестирования S-блока MCSSHA-3: *a* – результаты расчета эргодичности; *б* – автокорреляционная функция. По оси абсцисс отложена величина  $L/2$  (*a*) и номер итерации *j* (*б*)

Из сравнения рис. П.1 и П.2 можно сделать вывод о преимуществах полноцикловых подстановок перед прочими подстановками с достаточно большой длиной цикла. Это видно из того, что точки отображения, генерируемого полноцикловой подстановкой, равномернее заполняют фазовое пространство. Очевидно, что уменьшение длины цикла ведет к снижению криптографических свойств S-блока и, как следствие, алгоритма хеширования в целом.

По характеру поведения кривой на рис. П.2, *a* можно заключить, что данная подстановка (MCSSHA-3) обладает хорошими размешивающими свойствами, поскольку кривая достаточно быстро выходит на значение соответствующей инвариантной меры  $\rho = 1$  (чего нельзя сказать о подстановке Abacus).

#### Список литературы

17. Иванов М.А. Криптографические методы защиты информации в компьютерных системах и сетях. М.: Кудиц-Образ, 2001.
18. ГОСТ Р 34.11-94. Информационная технология. Криптографическая защита информации. Функция хеширования.
19. [http://csrc.nist.gov/groups/ST/hash/sha-3/Round1/submit\\_ons\\_rnd1.html](http://csrc.nist.gov/groups/ST/hash/sha-3/Round1/submit_ons_rnd1.html)
20. Шустер Г. Детерминированный хаос / Пер. с англ.; Под ред. А.В. Гапонова-Грехова, М.И. Рабиновича. М.: Мир, 1988.

УДК 530.182:577.3

## ОБ ОБУСЛОВЛЕННОСТИ НЕПЕРИОДИЧЕСКИХ АВТОКОЛЕБАНИЙ ПРОКСИМАЛЬНОГО ДАВЛЕНИЯ В НЕФРОНЕ НЕЛИНЕЙНЫМ ОТКЛИКОМ СОКРАТИТЕЛЬНОГО МЕХАНИЗМА АФФЕРЕНТНОЙ АРТЕРИОЛЫ

О.Н. Павлова, Д.Э. Постнов

Саратовский государственный университет  
E-mail: pavlova\_olga@yahoo.ru

Методами нелинейной динамики и анализа временных рядов исследуется вопрос о механизме возникновения хаотических автоколебаний, экспериментально наблюдаемых в физиологических экспериментах по исследованию процесса авторегуляции почечного кровотока в нефронах почки млекопитающих. Полученные результаты говорят в пользу гипотезы, согласно которой сложные колебательные режимы возникают за счет нелинейных характеристик сократительного механизма малого кровеносного сосуда (афферентной артериолы) на входе в нефрон.

**Ключевые слова:** нефрон, автоколебания, хаос, артериола, гипертензия, вейвлет.



### About Conditionality of Nonlinear Response of Miogenic Response of Afferent Arteriola for Irregular Self-Sustained Oscillations of Nephron Proximal Pressure

O.N. Pavlova, D.E. Postnov

By means of nonlinear dynamics and time series analysis we investigate the possible mechanisms for the onset of chaotic self-sustained dynamics in nephron tubular pressure that is observed experimentally. Our results suggests that the miogenic constriction mechanism of afferent arteriola plays the key role providing the nonlinear response on temporal variation of filtration rate.

**Key words:** nephron, self-oscillation, chaos, hypertension, wavelet.

Одним из важных достижений в области ренальной физиологии было открытие автоколебательной динамики в функционировании малых структурных элементов почки млекопитающих – нефронов.

Почка человека содержит около миллиона нефронов, каждый из которых способен самостоятельно выполнять ее главную функцию: регулировать химический и ионный состав плазмы крови посредством ее фильтрации [1, 2]. Каждый нефрон (рис. 1) содержит так называемый клубочек, играющий роль фильтра, с кровеносными сосудами, приносящими и отводящими поток крови (афферентная и эфферентная артериолы). Выделенный фильтрат проходит процедуру регуляции химического состава в процессе продвижения по проксимальному извитому каналцу и петле Генле. В месте перехода в дистальный каналец, петля Генле примыкает непосредственно к клубочку (т.н. *macula densa*). Здесь в зависимости от химического состава фильтрата меняется тонус клеток гладкой мускулатуры, окружающих стенку афферентной артериолы, и ее диаметр изменяется: артериола сжимается при чрезмерно интенсивной фильтрации и расслабляется при ее недостаточной интенсивности. Согласно результатам измерений, высокая концентрация ионов хлора соответствует минимуму давления в каналце [3].

Такой механизм, наряду с заметным временем прохождения фильтратом петли Генле (порядка 13 с), создает предпосылки для автоколебательной динамики. И действительно, в экспериментах на нефронах крысы *in vivo* P.P. Leysec и N.H. Holstein-Rathlow наблюдали автоколебания проксимального гидростатического давления с периодом 20–40 с [4, 5]. Была экспериментально доказана важная роль упомянутых выше факторов: каналцево-гломерулярной обратной связи (КГОС), наличия существенной задержки при продвижении фильтрата по петле Генле, а также миогенного отклика афферентной артериолы [6–8].

Как выяснилось, характеристики колебаний проксимального давления существенно зависят от состояния организма в целом.

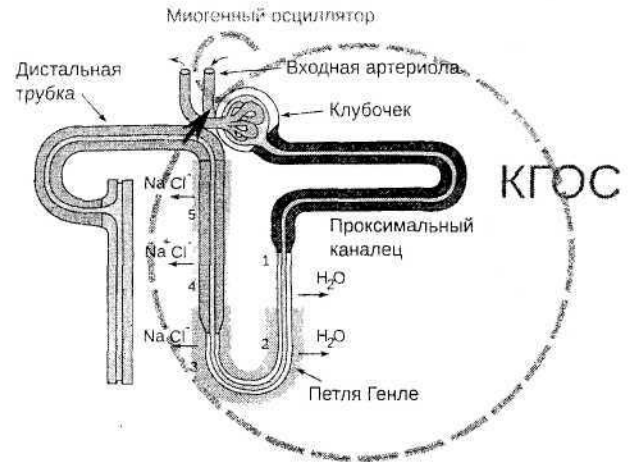


Рис. 1. Схематичное представление основных частей нефрона и путей их взаимодействия. По мере продвижения по проксимальному каналцу и петле Генле химический состав фильтрата меняется: в сегментах 1 и 2 удаляется вода, в сегментах 3, 4, 5 регулируется концентрация ионов натрия и хлора. Пунктирными окружностями обозначены обратные связи: каналцево-гломерулярная (КГОС) и внутренние связи миогенного осциллятора

Для крыс с повышенным артериальным давлением было экспериментально зафиксировано наличие нерегулярных (хаотических) колебаний проксимального давления [9]. На рис. 2 показано, как выглядит типичная временная реализация проксимального давления в случае нормы (а) и для гипертензивной крысы (б). Несмотря на неизбежную «загрязненность» сигнала шумами, различие очевидно: во втором случае амплитуда колебаний меняется быстро и на большую величину.

Такие различия в процессе авторегуляции почечного кровотока в норме и при гипертензии не могли не вызвать интереса исследователей как с точки зрения создания новых методов диагностики, так и в целях лучшего понимания механизмов работы ансамблей нефронов в почке. Однако однозначного объяснения того, как и почему возникают хаотические автоколебания при повышении артериального давления, на настоящий момент не предложено. Одна из причин этого в том, что эксперименты проводятся на цельной почке, без разрушения многообразных связей внутри и вне нефрона, а потому

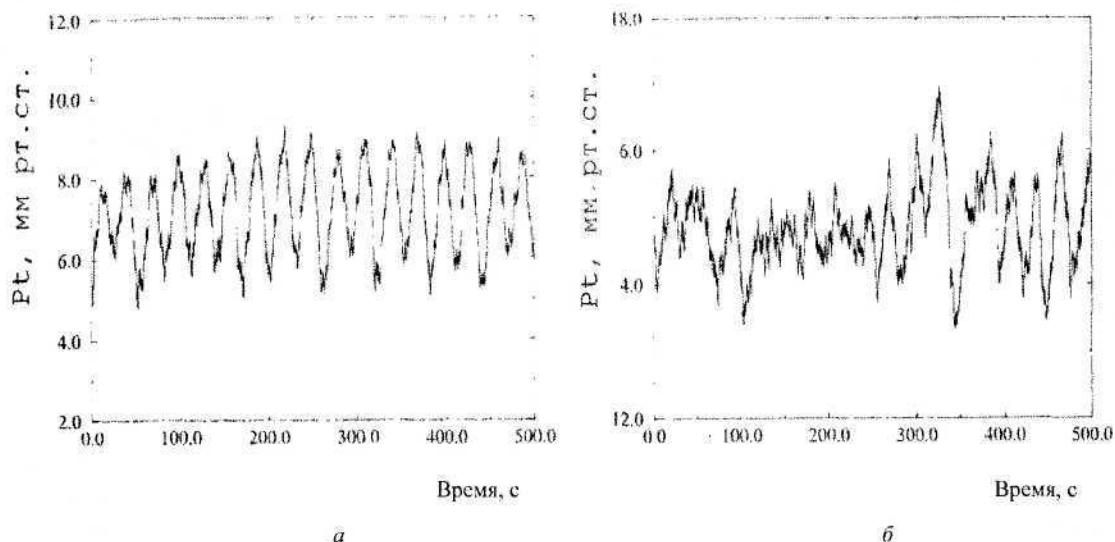


Рис. 2. Типичная временная реализация давления фильтрата в проксимальных канальцах нефронов нормотензивной (а) и гипертензивной (б) крысы

сложно утверждать, какие именно факторы отвечают за хаотизацию колебаний. Согласно гипотезе D.J. Marsh (США) принципиальную роль может играть взаимодействие нефрона с ближайшими соседями. Согласно другим данным, наличие хаотического режима коррелирует с чувствительностью клеток *macula densa* (по сути, с коэффициентом обратной связи) в цепи КГОС.

В данной работе мы анализируем эффект хаотизации колебаний в одиночном нефроне, используя методы и подходы, принятые в нелинейной динамике, а именно из экспериментальных данных мы реконструируем динамику радиуса афферентной артериолы и соответствующие фазовые проекции. Определенные выводы о механизме хаотизации колебаний позволяют сделать модификация математической модели и соответствующий вычислительный эксперимент. Полученные нами результаты работают в поддержку гипотезы, согласно которой хаос возникает вследствие сложного отклика механизма регуляции радиуса артериолы на переменное во времени внешнее воздействие.

#### 1. Экспериментально измеренные характеристики колебаний радиуса афферентной артериолы

На первом этапе исследований решалась задача извлечения информации о характере изменения сечения афферентной артериолы

из доступных экспериментальных записей временной реализации проксимального давления. Задача заключалась в том, чтобы выяснить, с каким типом поведения мы имеем дело – являются ли наблюдаемые процессы отражением автоколебательной динамики или нет? Если в отношении ритма КГОС сомнений в автоколебательной природе не возникало изначально, после того как соответствующие колебания были обнаружены и изучены в работах [10–13], то с миогенным ритмом ситуация иная. Длительное время считалось, что миогенная динамика является пассивным механизмом. К тому же математическая модель, предложенная в работах [14–16], описывает динамику афферентной артериолы уравнениями нелинейного осциллятора с затуханием. Однако после того как появилась возможность пронаблюдать временные зависимости мгновенных частот обоих ритмов авторегуляции, стали возникать сомнения в справедливости гипотезы о пассивности этого механизма [17, 18]. Тем не менее до настоящего времени детального исследования динамики артериол по экспериментальным данным не проводилось.

С целью анализа колебаний радиуса приносящей артериолы проводилось извлечение мгновенной амплитуды миогенного ритма в соответствии со следующей методикой, основанной на стандартном методе по-





строения «хребтов» (ridges) вейвлет-преобразования  $W(f, b)$  [19].

Хребты представляют собой линии локальных максимумов поверхности плотности энергии коэффициентов вейвлет-преобразования  $E(f, b) = |W(f, b)|^2$ , их расположение по параметру частоты  $f$  определяется для каждого значения параметра смещения (время)  $b$ . Если исключить из рассмотрения точки, соответствующие очень малым значениям  $|W(f, b)|$ , которые на практике могут быть связаны с зашумленностью сигнала, погрешностями вычислений и т.д. (порог задается с учетом особенностей анализируемого сигнала и уровня шума), то полученные хребты, изображенные в виде линий на плоскости время–частота ( $b, f$ ), соответствуют мгновенным частотам ритмов колебаний, присутствующих в анализируемом сигнале.

Такой способ анализа нестационарной многочастотной динамики обеспечивает возможность отслеживать временную эволюцию каждого ритма колебаний, представляющего

интерес для исследователя. Применительно к миогенным колебаниям это – ритм с частотой около 0.15 Гц. Поскольку каждая точка хребта ассоциируется с локальным энергетическим спектром, нахождение мгновенных частот колебательных процессов одновременно обеспечивает и определение мгновенной амплитуды  $a(t)$  этого процесса, которая отражает изменение радиуса приносящей артерии.

Для получения более наглядного представления о динамике колебательных процессов в нефроне целесообразно отметить также моменты времени, соответствующие локальным максимумам колебаний, связанных с механизмом КГОС и миогенным ритмом. Эти моменты нетрудно определить по локальным максимумам действительной части коэффициентов вейвлет-преобразования  $Re[W(f, b)]$ .

На рис. 3, а, в представлены типичные примеры вычисленных вышеописанным образом временных зависимостей мгновенной

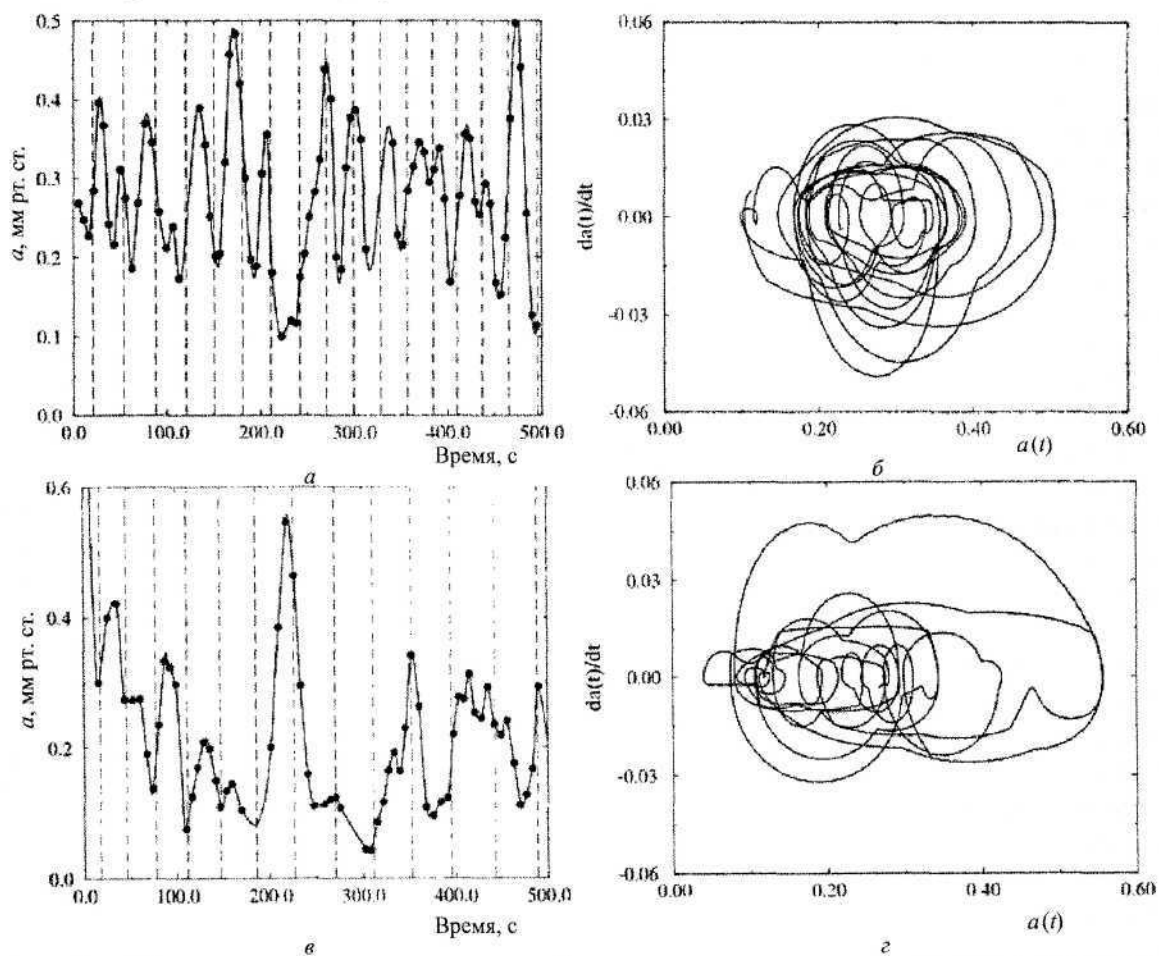


Рис. 3. Типичные примеры динамики мгновенной амплитуды миогенного ритма (а, в) и реконструированные фазовые портреты (б, г) двух гипертензивных крыс



амплитуды  $a(t)$  миогенного ритма нефронов гипертензивных крыс (сплошные линии). Черными точками отмечены времена локальных максимумов колебаний в артериоле (миогенного ритма), вертикальными пунктирными линиями – локальные максимумы колебательного процесса, связанного с механизмом КГОС. В данном случае мы ограничиваемся двумя примерами, поскольку полученные результаты визуально достаточно похожи для всех экспериментальных данных. Обращают на себя внимание несколько обстоятельств. Во-первых, на один период медленного ритма приходится разное количество периодов быстрого ритма (обычно 4–5), и это количество варьируется во времени, то есть не наблюдается длительных участков синхронного поведения, которое характеризовалось бы постоянным значением отношения частот колебаний.

Во-вторых, после очередного локального максимума медленного ритма динамика артериол ведет себя по-разному: радиус артериолы может увеличиваться, уменьшаться или демонстрировать слабые изменения во времени. Фактически получается, что миогенная динамика «живет собственной жизнью» и не является жестко привязанной к механизму КГОС. Это наблюдение опровергает ранее существовавшее мнение о том, что миогенная динамика является пассивным механизмом и может рассматриваться как осциллятор с затуханием.

Для исследования динамики артериол в терминах фазового пространства можно воспользоваться методом реконструкции [20], например, методом последовательного дифференцирования координат, и получить фазовые портреты, изображенные на рис. 3, б, г.

Они свидетельствуют о сложных нерегулярных колебаниях радиуса артериолы. Формальное применение метода расчета старшего показателя Ляпунова позволяет получить положительное значение порядка 0.1. Отметим, однако, что проведение расчетов ляпуновских показателей осложняется нестационарностью процессов, поэтому, строго говоря, не может служить достоверным индикатором хаотического режима для реальных экспериментальных данных. Тем не ме-

нее очевидно, что анализируемый режим колебаний радиуса артериолы является действительно сложным и нерегулярным.

## 2. Динамика математической модели одиночного нефрона при разомкнутой цепи КГОС и периодическом внешнем воздействии

Описанные выше результаты анализа экспериментальных данных ставят вопрос об адекватности принятого модельного описания динамики нефрона.

В настоящее время наиболее популярной моделью одиночного нефрона, хорошо описывающей как количественные характеристики процесса авторегуляции, так и их изменение в зависимости от управляющих параметров можно считать модель Барфреда [15]. Согласно ей, процесс авторегуляции отдельного нефрона может быть описан системой из шести дифференциальных уравнений:

$$\begin{aligned} \frac{dP_t}{dt} &= \frac{1}{C_{tub}} [F_{filt}(P_t, r) - F_{reab} - F_{Hen}(P_t)], \\ \frac{dr}{dt} &= v_r, \\ \frac{dv_r}{dt} &= \frac{P_{av}(P_t, r) - P_{eq}(r, \Psi(X_3, \alpha), T)}{\omega} - v_r d, \\ \frac{dX_1}{dt} &= F_{Hen}(P_t) - \frac{3X_1}{T}, \\ \frac{dX_2}{dt} &= \frac{3(X_1 - X_2)}{T}, \\ \frac{dX_3}{dt} &= \frac{3(X_2 - X_3)}{T}. \end{aligned}$$

Здесь  $P_t$  – гидростатическое давление в проксимальном канальце нефрона,  $F_{filt}$  – поток фильтрата из клубочка,  $F_{reab}$  – реабсорбируемый поток в проксимальном канальце,  $F_{Hen}$  – поток в начальной части петли Генли; коэффициент  $C_{tub}$  характеризует, насколько изменяется объем проксимального канальца при изменении давления;  $r$  – радиус приносящей артериолы, параметр  $d$  определяет скорость затухания колебаний (величину демпфирования),  $\omega$  есть отношение масса-эластичность стенок артериолы,  $v_r$  – скорость сокращения или растяжения артериолы,  $P_{eq}$  – равновесное давление в активной части приносящей артериолы,  $P_{av}$  – среднее давление в



активной части выносящей артериолы,  $X_1, X_2, X_3$  – вспомогательные переменные модели, а  $T$  – время задержки. Мускульный тон  $\psi$  есть функция, зависящая от скорости потока в дистальном канальце  $3X_3/T$  и параметра  $\alpha$ , определяющего петлевое усиление в цепи КГОС:

$$\psi\left(\frac{3X_3}{T}\right) = \psi_{\max} - \frac{\psi_{\max} - \psi_{\min}}{1 + e^{\alpha\left(\frac{3X_3}{T} - S\right)}},$$

$$S = 1 - \frac{1}{\alpha} \ln\left(\frac{\psi_{eq} - \psi_{\min}}{\psi_{\max} - \psi_{eq}}\right),$$

где  $\psi_{\min}, \psi_{\max}$  и  $\psi_{eq}$  – минимальный, максимальный и равновесный уровни активности,  $\frac{3X_3}{TF_{Hen0}}$  – выходящий поток из петли Генли, нормированный на равновесное значение потока  $F_{Hen0}$ . Более детальное обоснование модели и описание нелинейных функций можно найти в [15].

Используя описанную выше модель одиночного нефрона, мы пытались выявить механизмы, ответственные за возникновение хаотической динамики. Отметим, что данную модель можно представить в виде двух подсистем: первые 3 уравнения описывают динамику давления жидкости в проксимальном канальце и свойства артериол, а следующие 3 уравнения – задержку в реакции системы на изменение величины потока в петле Генли. Эти две подсистемы являются взаимосвязанными: переменная  $P_t$  первой из них входит в уравнение для  $X_1$  второй подсистемы, а переменная цепи задержки  $X_3$  влияет на динамику артериолы, входя в нелинейную функцию  $P_{eq}$ . Таким образом, мы имеем дело со взаимной связью двух подсистем, и в этом случае затруднительно определить, чем обусловлено наличие хаотической динамики – свойствами артериол или наличием временной задержки?

В поисках ответа на этот вопрос был проведен следующий модельно-вычислительный эксперимент: взаимная связь указанных выше подсистем заменялась на одностороннюю, когда свойства артериол и динамика в канальцах не оказывают влияния на

процессы в петле Генли. Для этого величина потока  $F_{Hen}$  в модели заменялась на заданное в явном виде периодическое внешнее воздействие вида  $F = B + A\cos(\Omega t)$ , где  $A$  и  $B$  характеризуют амплитуду воздействия и среднее значение потока соответственно, а  $\Omega$  – частоту воздействия.

Уравнения модели с заменой  $F_{Hen}$  на внешнее периодическое воздействие по смыслу соответствуют известной методике физиологического эксперимента [1], при которой проксимальный каналец перекрывается парафиновой пробкой, его содержимое отводится в микротрубочку, а поток в петлю Генле создается и поддерживается искусственно, с помощью другого стеклянного капилляра. Таким образом, канальцево-гломерулярная обратная связь оказывается разомкнутой, и по этой причине методика в англоязычной литературе носит название «open-loop». По аналогии мы будем обозначать математическую модель с заменой  $F_{Hen}$  на внешнее периодическое воздействие как «OL-модель».

Для оценки правомерности такого вычислительного эксперимента следует учесть тот факт, что петля Генли играет роль своеобразного фильтра нижних частот, не допуская быстрые вариации концентрации ионов хлора в районе клеток *macula densa*. На практике по результатам, опубликованным в [1], можно заключить, что колебания концентрации ионов хлора вблизи *macula densa* не содержат частотных составляющих миогенной динамики. Таким образом, можно считать, что периодическое внешнее воздействие имитирует автоколебательный режим исходной модели нефрона.

Если при взаимном влиянии двух подсистем при изменении параметров наблюдалась хаотическая динамика, то будет ли она и в случае разрыва одной связи? Проведем сопоставление режимов функционирования исходной модели и модели с разомкнутой КГОС. Оценка значений параметров периодического внешнего воздействия проводилась путем аппроксимации временной зависимости  $F_{Hen}$  в исходной модели Барфреда, при этом для периодического режима были получены следующие значения параметров:  $B = 0.2, A = 0.08, \Omega = 0.18$ .



Случаю нормального артериального давления соответствует значения параметров исходной модели Барфреда  $a = 12.0$  и  $T = 16$  с. При этом наблюдается периодический автоколебательный режим с периодом 30–40 с. Амплитуда колебаний давления составляет около 0.4 кПа, что также соответствует экспериментальным данным. OL-модель демонстрирует сходный режим функционирования при чуть большем среднем значении  $P_t$ . Таким образом, разорвав цепь КГОС, но подобрав подходящие параметры периодического внешнего воздействия, мы практически не нарушили работы первых трех уравнений модели.

Для крыс с повышенным артериальным давлением значения  $a$  больше, чем для крыс с нормальным давлением, поэтому математическое моделирование режима гипертонии может производиться путем увеличения параметра усиления КГОС. На рис. 4, а представлен пример хаотического поведения мо-

дели Барфреда, который наблюдается при значениях  $a = 32.0$  и  $T = 16$  с. В этом режиме старший ляпуновский показатель положителен, что доказывает наличие хаотического режима. Соответствующий фазовый портрет изображен на рис. 4, б. Здесь явно различимы быстрые колебания радиуса артериолы, промодулированные медленными колебаниями, обусловленными механизмом КГОС. Если OL-модель исследовать при таких же значениях параметров, то наблюдается сходное поведение: на рис. 4, в, г представлены временная реализация и проекция фазового портрета. Полного совпадения не наблюдается, во временной реализации более выражены быстрые колебания, фазовая проекция также несколько отличается. Однако различие не носит качественного характера. Можно утверждать, что при данных значениях управляющих параметров OL-модель воспроизводит хаотическое поведение, типичное для гипертонивных крыс.

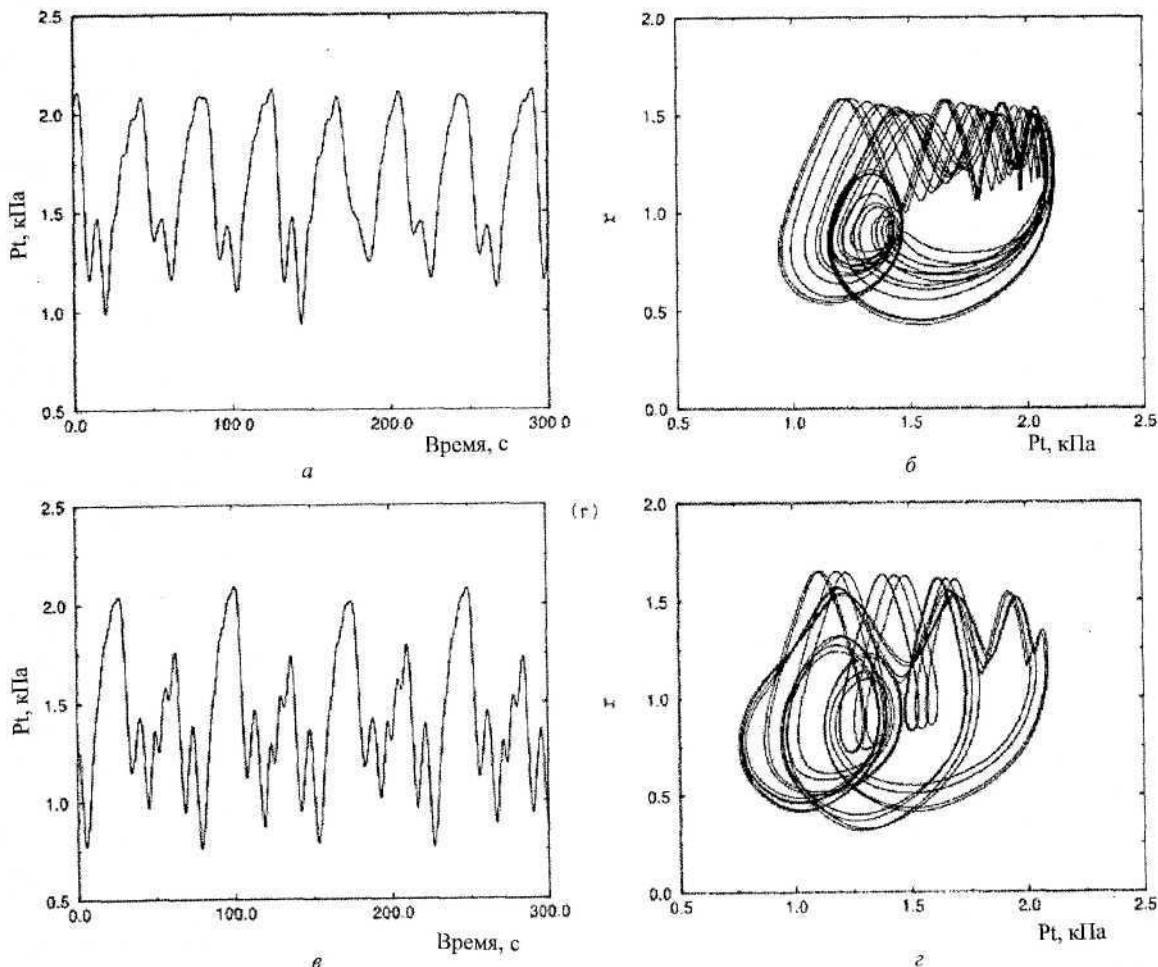


Рис. 4. Временные реализации (слева) и фазовые проекции (справа) для хаотического режима модели Барфреда (а, б) и модели с разомкнутой КГОС (в, г) при  $a = 32.0$  и  $T = 16$  с



Полученные результаты свидетельствуют о том, что возникновение хаотических режимов, наблюдаемых при повышенном артериальном давлении, определяется свойствами артериол, а не динамикой в петле Генли. В рамках исследования структурной устойчивости такого режима по отношению к выбору значений управляющих параметров был проведен бифуркационный анализ модели одиночного нефрона при изменении параметров  $d$  (параметр затухания),  $C_{tub}$  (эластичная растяжимость проксимального канальца) и  $a$  (усиление в цепи КГОС). Полученные диаграммы режимов в упрощенном виде приведены на рис. 5, а, в. Они, в частности, свидетельствуют о том, что переход к хаосу происходит при уменьшении параметра затухания колебаний в артериоле и в некотором

оптимальном диапазоне значений параметров  $C_{tub}$  и  $a$ . Для OL-модели, как видно из бифуркационных диаграмм, приведенных на рис. 5, б, г, хаотическая динамика наблюдается при сопоставимых значениях  $d$ . Вариация параметров  $\Omega$  и  $A$  качественно соответствует изменению  $C_{tub}$  и  $a$ : первый из них определяет частоту колебаний потока в петле Генли, а второй — силу воздействия. Как можно видеть, бифуркационные диаграммы для исходной и модифицированной модели качественно соответствуют друг другу. Из рис. 5, б следует, что хаотическая динамика в OL-модели наблюдается в полосе частот внешнего воздействия, соответствующей типичным значениям частоты собственных колебаний в модели Барфреда.

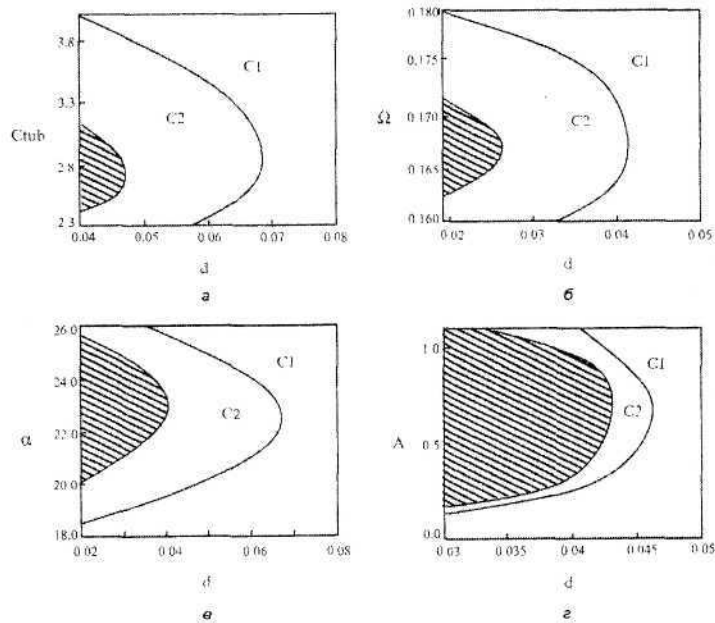


Рис. 5. Бифуркационные диаграммы, демонстрирующие переход к хаосу в модели Барфреда для режима колебаний с соотношением частот медленной и быстрой моды 1:4 (а, в) и OL-модели (б, г). Отображены области квазигармонических колебаний (C1), колебаний удвоенного периода (C2) и хаотической динамики (заштрихованы)

### 3. Обсуждение результатов и выводы

Описанные в предыдущих разделах результаты характеризуют динамику изменения радиуса афферентной артериолы одиночного нефрона как сложный процесс и могут оцениваться как с позиций соответствия физиологическим данным, так и с точки зрения

адекватности используемых в настоящее время математических моделей. В них «источником» хаотизации колебаний проксимального давления является именно механизм регуляции просвета афферентной артериолы посредством изменения тонуса клеток гладкой мускулатуры. В самом деле, выявленная



нами в сигнале давления в проксимальном канальце сложная динамика миогенного ритма практически отсутствует в потоке на выходе из петли Генле, а значит, привносится именно артериолой.

При обработке данных эксперимента невозможно установить, не привносится ли эта нерегулярная модуляция амплитуды миогенного ритма извне как влияние колебательной активности других близкорасположенных нефронов. Именно на проверке потенциальной возможности генерации хаотических колебаний в сократительном механизме артериолы и основан описанный в разделе 2 вычислительный эксперимент. Он основан на том, что модель Барфреда, несмотря на относительную простоту, разрабатывалась на основе информации о конкретных физиологических механизмах. А именно, величины  $P_{av}$  и  $P_{eq}$ , описываемые сложными математическими выражениями [15], отражают особенности строения эластичного сегмента артериолы, а выражение для величины  $\psi$  получено путем аппроксимации экспериментальных данных. Таким образом, в модели Барфреда «запрятана» дополнительная информация о свойствах миогенного механизма артериолы, которую мы и извлекаем с помощью специального вычислительного эксперимента.

Естественно, любая модель имеет свои границы применимости и не способна «ухватить» все особенности функционирования живых систем. В частности, известны и альтернативные модели одиночного нефрона. Попытки построения более детального описания процесса авторегуляции в нефроне предпринимаются в работе D.J. Marsh с соавторами [21]. Однако предложенная ими модель, хоть и является более детальной с точки зрения физиологии, не позволяет описать процесс хаотизации колебаний нефрона при повышении артериального давления и не демонстрирует переход к хаосу через каскад бифуркаций удвоения периода, в то время как экспериментальные исследования показывают наличие субгармоник в сигналах нефронов гипертензивных крыс, а это, в свою очередь, позволяет предположить наличие нелинейных механизмов перехода к хаосу (субгармонический каскад).

С точки зрения теории колебаний мы редуцировали исходную шестимерную автономную автоколебательную систему до трехмерной модели неавтономного нелинейного осциллятора (три последние уравнения в случае OL-модели линейны и играют роль «линии передачи» для периодического сигнала воздействия) и показали, что хаотическая динамика при этом сохраняется.

Итак, в данной работе методами нелинейной динамики и анализа временных рядов мы исследовали вопрос о возможном механизме возникновения хаотических автоколебаний, экспериментально наблюдаемых в физиологических экспериментах по исследованию почечного кровотока. Полученные результаты говорят в пользу гипотезы, согласно которой сложные колебательные режимы возникают за счет нелинейных характеристик сократительного механизма афферентной артериолы.

Авторы благодарят коллег О. Sosnovtseva, E. Mosekilde (Denmark Technical University), а также А.Н. Павлова (Саратовский госуниверситет) за полезные дискуссии в ходе выполнения работы.

*Работа выполнена при финансовой поддержке Федерального агентства по образованию в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 годы (государственный контракт № П374).*

#### Список литературы

1. Фундаментальная и клиническая физиология: Учебник для студ. высш. учеб. заведений / Под ред. А.Г. Камкина, А.А. Каменского. М.: Издат. центр «Академия», 2004. 1072 с.
2. Despopoulos A., Silbernagel S. Color atlas of physiology. Stuttgart: Georg Thieme Verlag, 1991.
3. Holstein-Rathlou N.-H., Marsh D.J. Oscillations of tubular pressure, flow and distal chloride concentration in rats // Amer. J. Physiol. 1989. Vol.256. P.1007–1014.
4. Holstein-Rathlou N.-H., Leyssac P.P. TGF-mediated oscillations in the proximal intratubular pressure: Differences between spontaneously hypertensive rats and Wistar-Kyoto rats // Acta Physiol. Scand. 1986. Vol.126. P.333–339.
5. Leyssac P.P., Holstein-Rathlou N.-H. Effects of various transport inhibitors on oscillating TGF pressure response in the rat // Pflügers Arch. 1986. Vol.407. P.285.
6. Holstein-Rathlou N.-H., Wagner A.W., Marsh D.J. Tubuloglomerular feedback dynamics and renal blood flow autoregulation in rats // Amer. J. Physiol. 1991. Vol.260. P.F53.



7. Chou K.H., Yu-Ming Chen, Mardarelis V.Z., Marsh D.J., Holstein-Rathlou N.-H. Detection of interaction between myogenic and TGF mechanisms using nonlinear analysis // Amer. J. Physiol. 1994. Vol.267. P.F160.
8. Feldberg R., Colding-Jorgensen M., Holstein-Rathlou N.-H. Analysis of interaction between TGF and the myogenic response in renal blood flow autoregulation // Amer. J. Physiol. 1995. Vol.269. P.F581.
9. Yip K.-P., Holstein-Rathlou N.-H., Marsh D.J. Chaos in blood flow control in genetic and renovascular hypertensive rats // Amer. J. Physiol. 1991. Vol.261. P.400–408.
10. Leyssac P.P., Holstein-Rathlou N.-H. Tubulo-glomerular feedback response: enhancement in adult spontaneously hypertensive rats and effects of anaesthetics // Pfluegers Arch. 1989. Vol.413. P.267–272.
11. Holstein-Rathlou N.-H., Marsh D.J. Oscillations of tubular pressure, flow, and distal chloride concentration in rats // Amer. J. Physiol. 1989. Vol.256. P.F1007–F1014.
12. Daniels F.H., Arendshorst W.J., Roberds R.G. Tubuloglomerular feedback and autoregulation in spontaneously hypertensive rats // Amer. J. Physiol. 1990. Vol.258. P.F1479–F1489.
13. Holstein-Rathlou N.-H., Wagner A.J., Marsh D.J. Tubuloglomerular feedback dynamics and renal blood flow autoregulation in rats // Amer. J. Physiol. 1991. Vol.260. P.F53–F68.
14. Holstein-Rathlou N.-H., Marsh D.J. A dynamic model of the tubuloglomerular feedback mechanism // Amer. J. Physiol. 1990. Vol.258. P.F1448–F1459.
15. Barfred M., Mosekilde E., Holstein-Rathlou N.-H. Bifurcation analysis of nephron pressure and flow regulation // Chaos. 1996. Vol.6. P.280–287.
16. Mosekilde E. Topics in Nonlinear Dynamics: Applications to Physics, Biology and Economic Systems. Singapore: World Scientific, 1996.
17. Sosnovtseva O.V., Pavlov A.N., Mosekilde E., Holstein-Rathlou N.-H. Bimodal oscillations in nephron autoregulation // Phys. Rev. E. 2002. Vol.66. P.061909.
18. Sosnovtseva O.V., Pavlov A.N., Mosekilde E., Yip K.-P., Holstein-Rathlou N.-H., Marsh D.J. Synchronization among mechanisms of renal autoregulation is reduced in hypertensive rats // Amer. J. Physiol. 2007. Vol.293. P.F1545–F1555.
19. Mallat S.G. A wavelet tour of signal processing. N.Y.: Academic Press, 1998.
20. Sauer T., Yorke J.A., Casdagli M. Embedology // J. Statistical Physics. 1991. Vol.65. P.579–616.
21. Marsh D.J., Sosnovtseva O.V., Chon K.H., Holstein-Rathlou N.-H. Nonlinear interactions in renal blood flow regulation // Amer. J. Physiol. 2005. Vol.288. P.R1143–R1159.

УДК 537.86:519.2

## ИДЕНТИФИКАЦИЯ ПОТЕНЦИАЛОВ ДЕЙСТВИЯ МАЛЫХ АНСАМБЛЕЙ НЕЙРОНОВ С ПРИМЕНЕНИЕМ ВЕЙВЛЕТ-АНАЛИЗА И МЕТОДА НЕЙРОННЫХ СЕТЕЙ

А.Н. Тупицын, А.И. Назимов, А.Н. Павлов

Саратовский государственный университет  
E-mail: tupitsin-anatoliy@yandex.ru

Обсуждается возможность решения задачи автоматической идентификации нейронных спайков во внеклеточных записях электрического потенциала на основе комбинированного алгоритма, предусматривающего применение нейронных сетей и дискретного вейвлет-преобразования. Иллюстрируется эффективность комбинированного подхода при анализе экспериментальных данных.

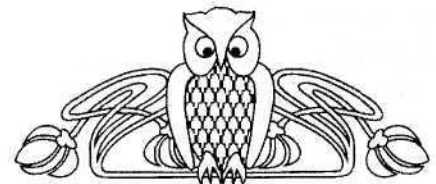
**Ключевые слова:** нейронные сети, вейвлет-анализ, сортировка спайков.

**Identification of Action Potentials of Small Neuron Ensembles Using Wavelet-Analysis and Neural Networks Method**

A.N. Tupitsyn, A.I. Nazimov, A.N. Pavlov

A possibility to solve the problem of automatic identification of neuronal spikes in the extracellularly recorded electrical potentials is discussed that is based on a combined approach assuming application of neural networks and the discrete wavelet-transform. Efficiency of the combined approach is illustrated in the analysis of experimental data.

**Key words:** neural networks, wavelet-analysis, spike sorting.



### Введение

Центральная нервная система (ЦНС) живых организмов непрерывно получает и обрабатывает сенсорную информацию, поступающую посредством взаимодействия с внешним миром. Исследование механизмов кодирования, представления и обработки данной информации ЦНС представляет собой одну из актуальнейших задач современного естествознания. Оптические, звуковые, тактильные и другие стимулы кодируются соответствующими рецепторами в последовательности электрических импульсов, которые передаются по нервным волокнам к «первым» нейронам, областям ЦНС, где осуществляется их первичная обработка. Далее



поступающая информация проходит еще несколько этапов обработки, прежде чем достигнуть коры головного мозга, где формируются соответствующие сенсорные образы. С каждым последующим этапом сложность экспериментальных исследований существенно усложняется. Хотя молекулярные и ионные механизмы, лежащие в основе кодирования информации, относительно хорошо изучены [1], менее понятны функциональные свойства последовательностей спайков как носителей информации – каким образом совокупность спайков отражает всю сложность и разнообразие окружающего мира?

Изучение процессов кодирования информации нейронными сетями в качестве процедуры предварительной обработки экспериментальных данных приводит к необходимости решения проблемы идентификации сигнала отдельной клетки в коллективной динамике ансамбля нейронов (или задачи сортировки спайков) [2–4]. Чтобы зарегистрировать электрическую активность нескольких нейронов и при этом не повредить их, традиционно проводятся эксперименты по внеклеточной записи электрического потенциала. Этот подход позволяет изучать динамику клеток, которые находятся в малой окрестности микроэлектрода, причем чем дальше будет находиться нейрон, тем меньше амплитуда сигнала, который принимается в точке регистрации.

К настоящему времени предложен ряд методов автоматической классификации спайков, включая методы, основанные на вейвлет-анализе [5–7], который считается одним из наиболее перспективных подходов к решению задачи распознавания образов. Однако следует отметить, что явные преимущества вейвлетных алгоритмов распознавания сигналов далеко не всегда очевидны [8–10], а уменьшение ошибки идентификации нейронных потенциалов действия при использовании специальных приемов [11] достигается за счет существенного увеличения времени вычислений. В связи с этим проблема разработки эффективных методов сортировки спайков по-прежнему остается весьма актуальной.

В данной работе предлагается комбинированный метод автоматического распознавания нейронных потенциалов действия, основанный на совместном применении вейвлет-анализа и техники искусственных нейронных сетей.

#### 1. Проблемы при решении задачи автоматической идентификации нейронных потенциалов действия

Традиционно считается, что каждый нейрон генерирует спайки одной и той же формы и амплитуды, а сигналы разных нейронов имеют хотя бы незначительные различия по форме. Однако задача идентификации усложняется из-за присутствия существенного фонового шума, обусловленного, в частности, динамикой удаленных клеток. На рис. 1 изображена довольно простая ситуация – наблюдаются хорошо отличающиеся сигналы двух типов, отмеченные стрелками (и их легко различить по амплитудам). В общем случае рассмотрение амплитуды как основной характеристики для идентификации сигналов разных нейронов оказывается неэффективным подходом [3]. Это связано с тем, что амплитуда зависит от расстояния между нейроном и микроэлектродом: если микроэлектрод располагается вблизи клетки, то сигнал, полученный от данного нейрона, будет превосходить сигналы отдаленных клеток и фоновый шум. Если же микроэлектрод будет расположен на одинаковом расстоянии от двух нейронов, то генерируемые ими сигналы будут похожими по амплитуде, поэтому их не удастся разделить выбором порогового уровня [2].

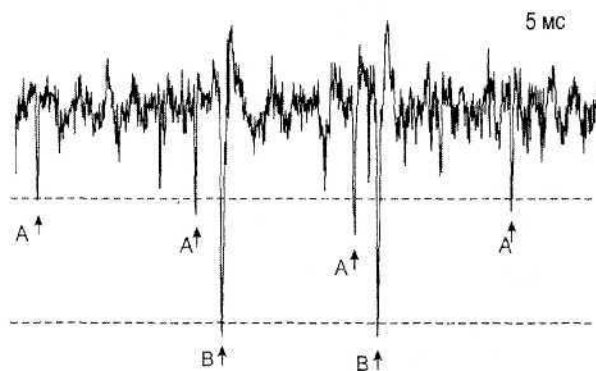


Рис. 1. Пример внеклеточной записи электрического потенциала, иллюстрирующий наличие спайков по крайней мере двух нейронов (обозначены стрелками А и В)





Основные проблемы при решении задачи автоматической идентификации нейронных потенциалов действия состоят в следующем. Во-первых, форма сигналов даже одного нейрона может варьироваться, и принципиальным моментом является, насколько эта вариабельность сопоставима с различиями между типичными сигналами двух соседних клеток, а стало быть, можно ли надежно их распознать? Известно, в частности, что в режиме «берстинга» в течение короткого промежутка времени одним нейроном генерируется серия импульсов, амплитуда которых уменьшается во времени. Во-вторых, часто происходит перекрытие спектров сигнала и шума, имеющего различную природу, и присутствие помех искажает формы спайков. Из-за перекрытия частотных диапазонов затрудняется фильтрация. Отмеченные факторы приводят к появлению ошибок распознавания сигналов. Например, было установлено [2], что при изучении проблемы кодирования информации в нейронных сетях экспериментатор, решающий задачу сортировки спайков вручную, может допускать ошибку классификации до 50% в случае одноканальной записи и до 10–30% при многоканальной записи (в зависимости от числа каналов). Такой способ сортировки (путем визуального анализа форм импульсов) крайне неэффективен, так как типичная экспериментальная запись может содержать более тысячи импульсов, и это обстоятельство приводит к невозможности тщательного разделения сигналов по группам. Из-за большого количества данных, которые подвергаются анализу, визуальный просмотр и разделение по группам, во-первых, займут достаточно большой промежуток времени, а во-вторых, беглый просмотр данных может привести к противоречивым решениям о принадлежности конкретного импульса той или иной группе. В свою очередь, это приведет к увеличению числа неправильно классифицированных сигналов, то есть к повышению ошибки идентификации. Таким образом, возникает необходимость автоматизации поиска решения данной задачи.

Проблема автоматической идентификации нейронных спайков неоднократно обсуждалась в научной печати [2–4]. Можно от-

метить несколько распространённых подходов, которые наряду с вышеупомянутым амплитудным детектированием включают анализ главных компонент (АГК) и вейвлет-анализ.

Анализ главных компонент [2] считается, возможно, самым эффективным среди классических методов автоматической идентификации спайков. Он представляет собой один из методов факторного анализа, главными задачами которого являются сокращение числа переменных (например, сокращение размерности динамических моделей), распознавание и сжатие образов, фильтрация от шума и т.д. Идея этого метода состоит в том, чтобы найти набор ортогональных векторов, которые характеризуют наиболее важные особенности формы сигналов. С точки зрения вычислительной математики задача сводится к поиску собственных векторов (главных компонент) ковариационной матрицы, построенной на основе экспериментальных данных. Для задачи идентификации, рассматриваемой в рамках данной работы, в качестве экспериментальных данных используются все нейронные спайки, отцентрированные по их максимумам или минимумам. При этом рассматриваются первые две главные компоненты (чтобы проанализировать проекцию на плоскость) или три главные компоненты (в случае трехмерного представления). Так как первые главные компоненты содержат максимальную информацию об анализируемых данных, такое проекционное представление можно использовать для решения задач идентификации сигналов. В качестве характеристик рассматриваются весовые (масштабные) коэффициенты, которые определяются как скалярное произведение анализируемого сигнала и соответствующей главной компоненты [2]. Недостатком данного метода является высокая чувствительность к низкочастотным флуктуациям [8].

Сравнительно недавно для решения задачи автоматической идентификации нейронных спайков стал широко применяться вейвлет-анализ [3, 4]. В основном использовалось дискретное преобразование, что объясняется быстродействием вычислений. В качестве характеристик для идентификации спайков в рамках данного подхода исполь-



зуются коэффициенты разложения по вейвлетам – солитоноподобным функциям [5–7]. Недостатком метода является проблема поиска «нужных» коэффициентов для идентификации сигналов из большого набора и отсутствие четких рекомендаций по их оптимальному выбору (выбор из «общих соображений» [3] часто оказывается неэффективен). Для устранения данного недостатка могут применяться комбинированные методы распознавания [9, 10] или специальные приемы уменьшения ошибки идентификации [11]. Однако более эффективным решением может служить применение вейвлетов в качестве инструмента предварительного анализа экспериментальных данных с привлечением искусственных нейронных сетей для более надежного разделения спайков по группам.

## 2. Применение нейронных сетей совместно с вейвлетами для решения задачи распознавания спайков

Нейросетевой метод идентификации сигналов [12–15] имеет свои особенности. В научной литературе существуют примеры готовых архитектур нейронных сетей, предназначенных для решения подобного рода задач. Но вместе с тем необходимо учитывать, что каждая конкретная задача имеет свои индивидуальные особенности, требующие внесения корректив в метод решения. Часто при решении сложных практических задач автономная нейронная сеть не может предоставить готовые решения. Поэтому в большинстве случаев нейронные сети интегрируют в многоступенчатые системы обработки данных, где сеть выполняет свою определенную роль для получения окончательного результата. Учитывая то обстоятельство, что необходимо различить похожие по форме сигналы в присутствии помех, необходим алгоритм, который, во-первых, проводит идентификацию сигналов с наибольшей чувствительностью, во-вторых, эффективен при наличии шума с различной статистикой, в-третьих, подстраивается под конкретные экспериментальные данные.

Для решения задачи идентификации спайков нервных клеток в данной работе предлагается алгоритм, объединяющий в себе методику вейвлет-анализа и технику ней-

ронных сетей. Вейвлет-анализ позволяет выявлять мелкие особенности в форме спайков и тем самым проводить их сортировку с максимальной чувствительностью, но данное качество может сыграть отрицательную роль при наличии существенного фонового шума в экспериментальных данных [9]. Поэтому вейвлет-анализ может служить прекрасным инструментом для поиска характерных особенностей в формах импульсов (то есть предварительной обработки) и передачи этой информации для дальнейшего анализа в нейронную сеть.

На втором этапе, располагая подобной информацией, будет применяться алгоритм, способный решать задачу более качественного распознавания образов. Если предположить, что на вход некоторой многослойной нейронной сети с прямым распространением [15] подаются данные, содержащие набор характерных признаков некоторых образов, то при наличии обучающей информации нейронная сеть будет в состоянии провести сортировку этих данных. Алгоритм обучения такой сети предполагает наличие обучающей выборки и эталонных значений, которые должны быть получены на выходе для этой выборки (так называемый алгоритм обучения с «учителем»). Иными словами, применительно к задаче идентификации спайков для успешной работы сети необходимо знать незашумленные формы импульсов, присутствующие в экспериментальной записи. А работа сети будет сводиться к сортировке спайков по группам.

Архитектура предлагаемого метода показана на рис. 2, а. Можно выделить несколько этапов обработки данных:

- 1) экспериментальные данные проходят обработку пороговым методом с целью выявления спайков, которые необходимо рассортировать;

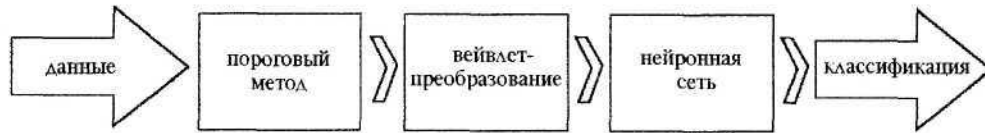
- 2) полученная последовательность подвергается дискретному вейвлет-преобразованию с помощью пирамидального алгоритма, в результате чего для каждой формы импульса (в проводимых расчетах она содержала 64 точки) получается набор вейвлет-коэффициентов; в качестве «материнского» вейвлета мы используем вейвлет Добеши;



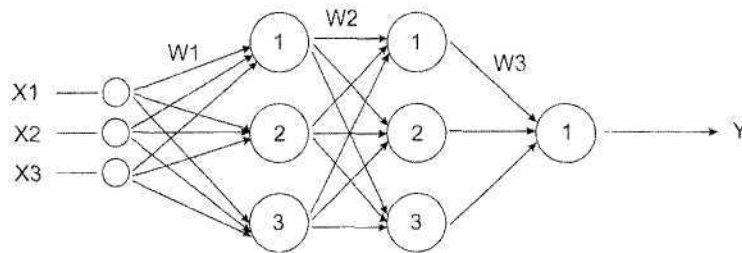
3) вейвлет-коэффициенты, характеризующие шумовую компоненту сигнала (то есть соответствующие малым масштабам), убираются, оставшиеся коэффициенты подаются на вход нейронной сети;

4) нейронная сеть с прямым распростра-

нением сигнала, содержащая три слоя, при получении на вход данных проводит сортировку спайков по группам и строит картину классификации в пространстве характеристики; в качестве функции активации сети используется логистическая функция.



а



б

Рис. 2. Метод, основанный на совместном применении вейвлет-анализа и техники нейронных сетей: а – структура метода; б – структура многослойной нейронной сети с прямым распространением сигнала

В большинстве случаев выбор архитектуры нейронной сети для решения конкретной задачи основывается на опыте разработчика, поэтому предложенный вариант является одним из множества возможных конфигураций. Однако необходимо учитывать следующие обстоятельства. Сеть с прямым распространением сигнала или многослойный перцептрон [14] является базовой и хорошо изученной структурой организации сети (рис. 2, б). Что касается выбора количества нейронных слоев и количества элементов в каждом из них, то он основывается на дилемме стабильности–пластичности [12]. Очевидно, что чем более сложной будет организация сети в плане общего количества элементов, тем выше будут адаптивные способности системы. Но следует заметить, что адаптивность не всегда ведет к устойчивости. Поэтому структуризация нейронной сети должна обеспечивать стабильность и гибкость ее работы одновременно.

### 3. Результаты

*3.1. Анализ тестовых данных.* Чтобы проверить эффективность предложенного метода, в частности, исследовать влияние статистики фонового шума на качество решения задачи автоматической идентификации сигналов с помощью данного алгоритма, проводилось его тестирование. Из экспериментальных данных (внеклеточные записи активности малых групп нейронов) выбирались две разные формы импульсов, наименее зашумленные (каждая из них содержала 64 отсчета). Соответствующие формы импульсов многократно повторялись, в результате был получен тестовый сигнал, состоящий из 1892 импульсов (по 946 импульсов разного типа). К полученной последовательности добавлялся цветной шум, полученный путем полосовой фильтрации нормально-распределенного случайного процесса. Полученный зашумленный сигнал анализировался на основе предложенного алгоритма идентификации.



Нейронная сеть включала три слоя, входной слой содержал 64 элемента (число, равное размеру входного вектора), второй слой содержал 32 элемента и выходной слой содержал 2 элемента. Сеть необходимо было обучить таким образом, чтобы при получении на входной слой вейвлет-коэффициентов разложения сигнала она осуществляла идентификацию спайка с наименьшей ошибкой и демонстрировала слабую чувствительность к влиянию статистики шума. Для этой цели использовался алгоритм обратного распространения ошибки [12].

В качестве обучающей выборки  $x_i$  сети подавались 100 образов в виде вейвлет-коэффициентов, соответствующих форме первого импульса и 100 образов, соответствующих форме второго импульса. Соответственно метод обучения содержал 100 эпох. Задача сети состояла в том, чтобы получить наиболее качественное разделение групп спайков в пространстве характеристик, поэтому в качестве эталонных выходных значений  $y_i$  задавались векторы с набором координат для первой и второй формы импульса в данном пространстве. Таким образом, получив на вход вектор, содержащий вейвлет-коэффициенты образа первого импульса, нейронная сеть должна ассоциировать его, например, с вектором координат  $[0,1; 0,1]$ , а для образа второго импульса – с вектором  $[0,5; 0,5]$ . После завершения такой процедуры нейронную сеть можно считать обученной. Для проверки работоспособности сети на ее вход целесообразно подать первоначальную последовательность, содержащую вейвлет-коэффициенты незашумленных спайков. Проведя такие вычисления, мы убедились, что метод позволяет качественно решать задачу идентификации разных форм потенциалов действия и, как минимум, не уступает по своим возможностям стандартным подходам.

### 3.2. Анализ экспериментальных данных.

При проведении тестовых экспериментов с искусственно сгенерированными данными известны незашумленные формы импульсов и количество групп идентифицируемых сигналов. Поэтому для обучения нейронной сети можно с успехом применять алгоритмы

обучения с «учителем», так как доступна обучающая выборка и эталонные данные, которые должны быть определены на выходе сети. На основе этих данных получается обученная нейронная сеть, способная решать задачу классификации импульсов в рамках конкретного численного эксперимента. Для реальных нейрофизиологических данных информация о форме незашумленных сигналов и о количестве разных видов спайков отсутствует. Главным барьером для использования предложенного совместного метода вейвлет-анализа и нейронных сетей в этом случае является обучение нейронной сети.

Исходя из вышеизложенных аргументов, имеет смысл несколько модифицировать данный метод, рассмотрев следующую последовательность действий:

1) использование порогового метода для выделения спайков, которые необходимо отсортировать;

2) проведение предварительной классификации методом анализа главных компонент или методом дискретного вейвлет-преобразования; получение картины кластеризации в пространстве характеристик;

3) нахождение областей с повышенной плотностью точек в полученных кластерах, получение форм спайков, соответствующих данным точкам (они будут ближе всего к незашумленным формам сигналов);

4) проведение вейвлет-преобразования применительно к выделенным импульсам, получение набора вейвлет-коэффициентов;

5) использование полученных вейвлет-коэффициентов в качестве обучающей выборки для нейронной сети;

6) проведение вейвлет-преобразования для исходного сигнала, анализ экспериментальных данных обученной сетью.

Таким образом, в качестве обучающей выборки для нейронной сети служат спайки, соответствующие центральной части разных кластеров, полученных другими методами автоматической идентификации, такими как анализ главных компонент или стандартный вейвлет-анализ. В качестве эталонных значений на выходе сети используются векторы координат пространства характеристик.



Как показывает практика, часто при использовании методов автоматического распознавания формы сигнала (АГК или вейвлет-анализ) получаемая картина точек на плоскости характеристик бывает неоднозначной. Кластеры на плоскости характеристик, соответствующие разнородным группам спайков, могут располагаться близко друг к другу, перекрываться, в результате чего точки из одного кластера ошибочно попадают в другой (рис. 3). Предложенный ме-

тод совместного использования вейвлет-анализа и нейронных сетей за счет индивидуального подхода к конкретной задаче позволяет уменьшить ошибку идентификации, то есть увеличить расстояние между кластерами на плоскости характеристик (см. рис. 3). Индивидуальность подхода достигается благодаря тому, что обучение нейронной сети происходит на выборке, содержащей спайки непосредственно из экспериментальных данных, представленных для анализа.

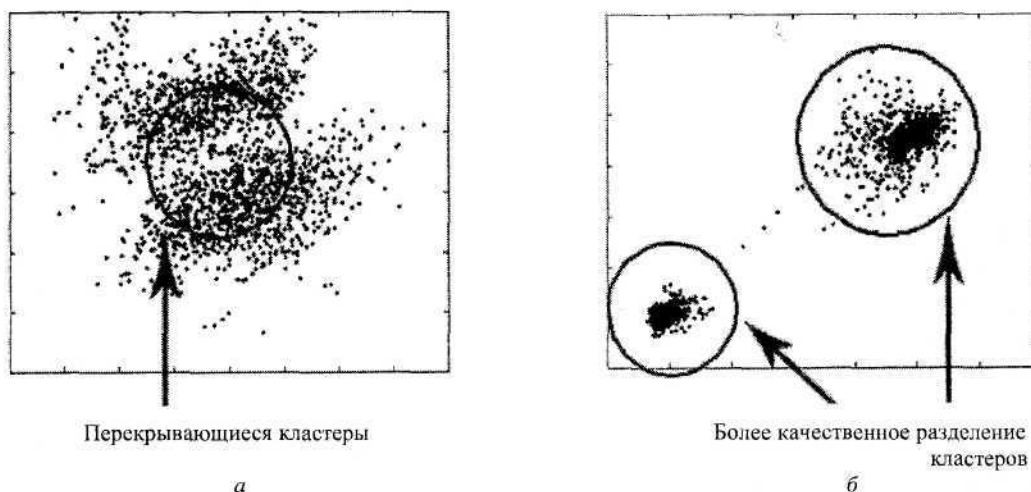


Рис. 3. Пример кластеризации методом вейвлет-анализа (*a*) и предложенным совместным методом (*б*)

Для того чтобы исследовать эффективность метода совместного использования вейвлет-анализа и нейронных сетей в реальных условиях, был проведен анализ нейрофизиологических данных предложенным алгоритмом. Данные снимались *in vivo* при помощи специального микроэлектрода с крысы. Микроэлектрод располагался в головной части животного в районе тройничного нерва. Для снятия необходимой сенсорной информации производилось периодическое механическое раздражение вибриссы воздушным потоком с некоторой фиксированной частотой. При проведении такого эксперимента микроэлектрод регистрировал активность некоторого нейронного ансамбля, то есть сигналы группы нейронов вблизи микроэлектрода. В результате получалась последовательность, содержащая набор разных форм спайков, которые для дальнейшего исследования генерируемого информационного

кода необходимо рассортировать. Полученные картины кластеризации при использовании предложенного метода представлены на рис. 4.

Как видно из рис. 4, использование стандартного вейвлет-анализа для задачи сортировки данных приводит к получению нечеткой картины кластеризации. Экспериментальная последовательность содержит три группы импульсов, но кластеры недостаточно хорошо отделены друг от друга и имеются общие точки, которые могут быть ошибочно отнесены не к той группе, к которой они на самом деле принадлежат. Следуя алгоритму предложенного метода, из центров получившихся кластеров авторы статьи выбирали по 50 точек, и данная выборка использовалась в качестве обучающей для нейронной сети. Произведя алгоритм обучения, на вход сети подавали ряд вейвлет-коэффициентов первоначальной последовательности импульсов.

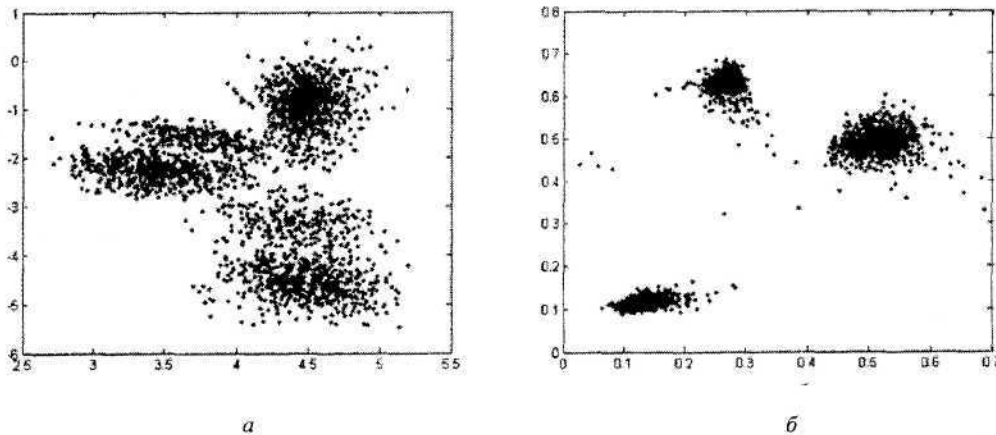


Рис. 4. Картина кластеризации при применении совместного метода вейвлет-анализа (а) и нейронных сетей на реальных экспериментальных данных (б)

Получившиеся картины кластеризации наглядно показывают улучшение решения задачи автоматического распознавания сигналов – более четкое разделение кластеров.

#### Заключение

В данной работе предложен метод идентификации спайков нервных клеток, основанный на совместном применении вейвлет-анализа и искусственных нейронных сетей. На первом шаге данного алгоритма осуществляется выявление мелких особенностей в формах спайков с использованием вейвлетов, на втором шаге эта информация передается в нейронную сеть для дальнейшего анализа и более качественного распознавания образов. Продемонстрировано, что данный подход позволяет эффективнее решать задачу автоматической сортировки нейронных спайков по сравнению с широко используемыми методиками (АГК, обычный вейвлет-анализ).

Предложенный метод, предусматривающий совместное применение вейвлет-анализа и техники нейронных сетей, является перспективным способом решения задач автоматического распознавания сигналов. Более того, возможности этого метода не ограничиваются только сигналами импульсного типа – с тем же успехом может решаться задача распознавания самых разных сигналов, поскольку форма сигнала не играет принципиальной роли с точки зрения методики анализа.

Авторы выражают искреннюю признательность В.А. Макарову за многочисленные дискуссии и предоставленные экспериментальные данные.

*Работа выполнена при финансовой поддержке Федерального агентства по образованию в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 годы (государственный контракт № П451).*

#### Список литературы

1. Tuckwell H.C. Introduction to theoretical neurobiology. Cambridge: Cambridge University Press, 1988.
2. Lewicki M. A review of methods for spike sorting: the detection and classification of neural potentials // Net. Com. Neu. Sys. 1998. Vol.9. P.R53–R78.
3. Letellier J., Weber P. Spike sorting based on discrete wavelet transform coefficients // J. of Neuroscience Methods. 2000. Vol.101. P.93–106.
4. Hulata E., Segev R., Ben-Jacob E. A method for spike sorting and detection based on wavelet packets and Shannon's mutual information // J. of Neuroscience Methods. 2002. Vol.117. P.1–12.
5. Daubechies I. Ten lectures on wavelets. Philadelphia: S.I.A.M., 1992.
6. Meyer Y. Wavelets: Algorithms and applications. Philadelphia: S.I.A.M., 1993.
7. Addison P.S. The illustrated wavelet transform handbook: applications in science, engineering, medicine and finance. Philadelphia: IOP Publishing, 2002.
8. Думский Д.В., Павлов А.Н., Тутицын А.Н., Макаров В.А. Классификация нейронных потенциалов действия на основе вейвлет-преобразования // Изв. вузов. Прикладная нелинейная динамика. 2005. Т.13, №5–6. С.77–98.
9. Pavlov A.N., Makarov V.A., Makarova I., Panetsos F. Separation of extracellular spikes: when wavelet based methods outperform the principle component analysis // Lecture Notes



in Computer Science / Eds. J. Mira, J.R. Alvarez. Berlin, 2005. P.123–132.

10. Pavlov A.N., Makarov V.A., Makarova I., Panetos F. Sorting of extracellular spikes: When wavelet based methods outperform the principle component analysis // Natural Computing. 2007. Vol.6. P.269–281.

11. Макаров В.А., Павлов А.Н., Тулицын А.Н. Сортировка нейронных спайков на основе параметрического вейвлет-анализа с адаптивной фильтрацией // Цифровая обработка сигналов. 2008. №3. С.26–31.

12. Haykin S. Neural networks. A comprehensive foundation. New Jersey: Prentice Hall, 1999.

13. Kohonen T. Selforganization and associative memory. N.Y.: Springer-Verlag, 1989.

14. Hopfield J., Tank D. Neural computation of decision in optimization problems // Biol. Cybernet. 1985. Vol.52. P.141–152.

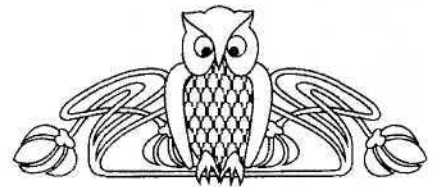
15. Callan R. The essence of neural networks. New Jersey: Prentice Hall, 1999.

УДК 621.371

## О ЗАКОНАХ СОХРАНЕНИЯ ЭНЕРГИИ И ИМПУЛЬСА ЭЛЕКТРОМАГНИТНОГО ПОЛЯ В СРЕДЕ И ПРИ ДИФРАКЦИИ ПЛОСКОЙ ВОЛНЫ НА ПРОВОДЯЩЕЙ ПЛАСТИНЕ

М.В. Давидович

Саратовский государственный университет  
E-mail: DavidovichMV@info.sgu.ru



На основе строгого нестационарного подхода получены общие нестационарные балансные уравнения для плотностей энергии и импульса системы поле–вещество, основанные на нестационарных определениях этих плотностей с зависимостью от предыстории процесса создания поля, а также определены скорости их переноса. Конкретное рассмотрение произведено для простейшего закона дисперсии, определяемого связанной с проводимостью диссипацией. Для него найдены плотность электромагнитной энергии в плоской монохроматической волне и скорости: фазовая, групповая, а также скорости переноса энергии и импульса. Показано, что плотность энергии на низких частотах в этом случае имеет вид электростатической плотности с заменой диэлектрической проницаемости на ее реальную часть, а скорость энергии в плоской монохроматической волне с указанной дисперсией равна фазовой скорости. Групповая же скорость может превышать скорость света. Показана необходимость использования формы Минковского для плотности импульса и определена скорость его переноса, которая в рассматриваемом случае также совпадает с фазовой скоростью. Продемонстрировано выполнение законов сохранения энергии и импульса в плоской электромагнитной волне в среде и при её дифракции на проводящей магнитодиэлектрической пластине.

**Ключевые слова:** тензор энергии-импульса, плотность импульса по Абрагаму и по Минковскому, скорость переноса энергии, скорость переноса импульса.

**On Conservation Laws for Energy and Momentum in Electromagnetic Field in Media and Under Plane Wave Diffraction on Conducting Plate**

M.V. Davidovich

The general nonstationary balance equations for energy and momentum densities of field-matter system based on rigorous nonstationary approach for their definitions with dependence from the field creation

prehistory have been obtained. Also the transport velocities of these densities have been derived. The concrete examination and detailed consideration have been performed for simplest dispersion law which is defined by the conductivity connected with the dissipation. There are following parameters which have been found for plane monochromatic wave under this law: the energy density, the phase velocity, the group velocity, and the transport velocities of energy and momentum. It has been shown that the energy density has the static form in which the dielectric permittivity must be replaced by its real part, and the energy transport velocity coincides with the phase velocity. The group velocity in this case may exceed the light velocity in the vacuum. It has been also shown that correct form of momentum density is the Minkowski one, and the momentum transport velocity in this case also coincides with phase velocity. The energy and momentum conservation have been shown for plane electromagnetic wave in the conducting medium and for plane wave diffraction on the conducting plate.

**Key words:** energy-momentum tensor, the Abraham and Minkowski momentum densities, energy transport velocity, momentum transport velocity.

### Введение

В электродинамике сплошных сред до настоящего времени сохраняется парадоксальное состояние, когда вот уже более ста лет нет однозначности в отношении выбора правильного определения тензора энергии-импульса (ТЭИ) [1–13]. Имеются два основных определения: Минковского [1] и Абрагама [2]. Соответственно имеется много публикаций как за определение Минковского, например, [13], и против определения Абра-



гама, так и наоборот – за Абрагама и против Минковского (см., например, публикации [3–6], последние обзоры [11, 12], работу [13] и литературу в них). Кроме этого имеются работы, авторы которых утверждают или считают, что указанные два определения эквивалентны (например, [3, 12]), но все-таки тензор Абрагама более предпочтительный или правильный. Тем не менее тензором Минковского тоже можно пользоваться, он часто более удобен и точнее соответствует сплошной среде [3]. В других публикациях тензор Минковского считается более верным. Есть ряд публикаций об экспериментальных подтверждениях и опровержениях как первого, так и второго из определений [12]. В частности, есть публикации по измерению силы Абрагама  $\vec{f}^A$  – величины, являющейся довеском к производной по времени от плотности импульса по Абрагаму  $\partial_i \vec{g}^A$ , в результате которого получается производная  $\partial_i \vec{g}^M$  для плотности импульса по Минковскому [3]. Заметим, что все указанные эксперименты произведены для квазистационарных или нестационарных (импульсных) процессов, а объемная сила Абрагама не эквивалентна сумме сил Лоренца, действующих на электрический и магнитный токи поляризации вещества [5]. Далее (во избежание недоразумений) под электромагнитным импульсом, или цугом, понимаем нестационарную электромагнитную волну, а под импульсом электромагнитного поля  $\vec{G}$  или моментом – интеграл по объему от плотности линейного импульса системы поле-вещество  $\vec{g}$ .

Упомянутая неоднозначность породила также ряд попыток определить и вывести ТЭИ по-другому, например, с использованием микроскопической электродинамики [7, 8], уравнений движения вещества или с использованием теоремы Нётер [13]. Считается, что однозначно можно определить только ТЭИ системы поле-вещество, а по отдельности указанные величины определяются неоднозначно. Между тем неоднозначность в определении плотности импульса поля ведет к неопределенности скорости его переноса

$\vec{v}_i^{(EM)}$ , скорости движения вещества  $\vec{v}^{(M)}$  и полной скорости переноса импульса поле-вещество  $\vec{v}_i$ . Также из нее следует неоднозначность силы давления электромагнитной волны на вещество.

В данной работе получены нестационарные балансные уравнения для энергии и импульса и введены нестационарные определения их плотностей, зависящие от предыстории процесса создания поля. На основе этого приведены выражения для скорости переноса энергии и импульса системы поле-вещество, а также для скорости переноса импульса поля и вещества. Для простоты общие полученные результаты конкретизируются рассмотрением одномерных задач: а) о плоской электромагнитной волне в среде с дисперсией, обусловленной проводимостью; б) о её дифракции на плоскопараллельной проводящей пластине. Указанные соотношения для монохроматического процесса позволяют построить ТЭИ и определить скорости переноса энергии и импульса системы поле-вещество, которые для указанной дисперсии совпадают с фазовой скоростью.

### 1. Балансные уравнения для энергии и импульса

Всегда, когда возникают парадоксы, необходимо искать, где произошла подмена понятий или неправомерно применены некие понятия к явлениям [14]. В случае ТЭИ, плотностей энергии и импульса в среде это заключается в неправомерной замене нестационарных понятий стационарными. В частности, плотность энергии поля в веществе определяют так:

$$u(\vec{r}, t) = [\vec{D}(\vec{r}, t) \vec{E}(\vec{r}, t) + \vec{B}(\vec{r}, t) \vec{H}(\vec{r}, t)] / 2,$$

т.е. как в статике, что неверно [14–16]. Аналогично плотность импульса по Абрагаму определена в виде  $\vec{g}^A = \vec{S} / c^2$ , а по Минковскому – в виде

$$\vec{g}^M = \vec{D}(\vec{r}, t) \times \vec{B}(\vec{r}, t) = n^2 \vec{S}(\vec{r}, t) / c^2 = n^2 \vec{g}^A,$$

где  $n = \sqrt{\epsilon \mu}$  – показатель преломления (индекс рефракции, или коэффициент замедления),  $\vec{S}(\vec{r}, t) = \vec{E}(\vec{r}, t) \times \vec{H}(\vec{r}, t)$  – вектор Пойн-





тинга. Соответственно в подавляющем большинстве работ, касающихся ТЭИ, полагается, что материальные соотношения имеют вид

$$\begin{aligned}\vec{D}(\vec{r}, t) &= \varepsilon_0 \varepsilon(\vec{r}) \vec{E}(\vec{r}, t), \\ \vec{B}(\vec{r}, t) &= \mu_0 \mu(\vec{r}) \vec{H}(\vec{r}, t),\end{aligned}$$

т.е. в предположении отсутствия временной (частотной) дисперсии, что также имеет место лишь в статике. Приведенная запись соответствует неоднородной среде. В большинстве работ  $\varepsilon$  и  $\mu$  – просто числовые константы. Мы рассмотрим материальные уравнения в обобщенной форме Казимира [17]:

$$\begin{aligned}\vec{D}(\vec{r}, t) &= \varepsilon_0 \partial_t^{-1} \partial_r^{-1} \left( \hat{\varepsilon}(\vec{r}, \vec{r}', t - t') \vec{E}(\vec{r}', t') \right), \\ \vec{B}(\vec{r}, t) &= \mu_0 \partial_t^{-1} \partial_r^{-1} \left( \hat{\mu}(\vec{r}, \vec{r}', t - t') \vec{H}(\vec{r}', t') \right),\end{aligned} \quad (1)$$

хотя есть и другие формы (например, Ландау–Лифшица) [18]. Здесь введены следующие интегральные операторы:

$$\begin{aligned}\partial_t^{-1}(f(t')) &= F(t) = \int_0^t f(t') dt', \\ \partial_r^{-1}(\varphi(\vec{r}, \vec{r}')) &= \Phi(\vec{r}) = \int_V \varphi(\vec{r}, \vec{r}') d^3 r',\end{aligned} \quad (2)$$

где под интегралами могут стоять как скалярные, так и векторные функции, а  $d^3 r' = dV'$  есть элемент объема точки источника. В первом интеграле нижний предел можно взять в виде  $-\infty$ , и тогда должно выполняться  $f(t) \rightarrow 0$  при  $t \rightarrow -\infty$ . Для простоты мы рассматриваем однородные во времени процессы, в силу чего ядра в (1) зависят от разности  $t - t'$ , т.е. выполняется принцип причинности:

$$\hat{\varepsilon}(\vec{r}, \vec{r}', t - t') = \hat{\mu}(\vec{r}, \vec{r}', t - t') = 0 \text{ при } t' > t.$$

Во втором интеграле объем выбирается из условия  $|\vec{r} - \vec{r}'|/c \leq t - t'$ , т.е. опять выполняется принцип причинности. Он означает, что вклад в индукции вносят поля в точках, расположенных не далее расстояния, за которое свет проходит путь  $c(t - t')$ , что соответствует учету пространственной дисперсии (здесь всегда  $t \geq t'$ ). Обычно пространственной дисперсии соответствует гораздо мень-

шая область, чем обозначенная выше. Ядра в (1) в общем случае тензорные и неоднородные по координатам, что соответствует анизотропным неоднородным средам.

Запишем уравнения Максвелла в общей форме:

$$\begin{aligned}\nabla \times \vec{H}(\vec{r}, t) &= \partial_t \vec{D}(\vec{r}, t) + \vec{J}^e(\vec{r}, t), \\ -\nabla \times \vec{E}(\vec{r}, t) &= \partial_t \vec{B}(\vec{r}, t) + \vec{J}^m(\vec{r}, t).\end{aligned} \quad (3)$$

Здесь, как обычно,  $\nabla \equiv \partial_r$  означает векторный дифференциальный оператор. Смысл уравнений Максвелла (3) достаточно прост: это полный баланс токов, причем в левой части стоит полная плотность тока (электрического в первом уравнении и магнитного во втором уравнении), а в правых частях стоят суммы соответствующих токов смещения и сторонних токов. Уравнения (3) наиболее полные, поскольку всё влияние среды (в частности, токи проводимости) учтено в материальных соотношениях. Учёт электрической проводимости в форме  $\vec{J}_\sigma^e(\vec{r}, t) = \sigma^e(\vec{r}) \vec{E}(\vec{r}, t)$  (т.е. так, как для постоянного тока) возможен путем следующего представления ядра:

$$\begin{aligned}\hat{\varepsilon}(\vec{r}, \vec{r}', t - t') &= \delta(t - t') \left[ \hat{\varepsilon}(\vec{r}, \vec{r}') + \right. \\ &\left. + \left( \sigma^e(\vec{r}) / \varepsilon_0 \right) \partial_t^{-1} \right] + \kappa^e(\vec{r}, \vec{r}', t - t'),\end{aligned}$$

где  $\kappa^e(\vec{r}, \vec{r}', t - t')$  – ядро оператора электрической восприимчивости,  $\partial_t \partial_t^{-1} = I$  – единичный оператор. В простом случае только частотной дисперсии такой учет проводимости соответствует появлению полюса у спектральной функции диэлектрической проницаемости на нулевой частоте [17]. В общем случае следует использовать формулу Друде для частотной зависимости проводимости, что соответствует виду ядра  $\hat{\varepsilon}$  интегрального оператора для плазмы. Во второе уравнение (3) входит сторонний магнитный ток. Хотя магнитный заряд (монополю Дирака) до сих пор не обнаружен, введение  $\vec{J}^m(\vec{r}, t)$  весьма полезно для симметрии, поскольку сторонние магнитные токи могут быть эквивалентны определенным конфигурациям сто-



ронных электрических токов. Далее мы считаем, что поле при  $t < 0$  отсутствовало. Соответственно плотность энергии и импульса поля и вещества (с точностью до энергии покоя) до момента  $t < 0$  были равны нулю. В момент  $t_0 = 0$  включаются источники, которые производят работу по созданию поля и изменению энергии и импульса поля и вещества. Обычно считается, что энергия источников носит неэлектромагнитный характер, что удобно чисто математически, хотя физически часто эта энергия все же электромагнитная, но действует вне объема рассмотрения поля. Часть произведенной энергии диссипируется в тепло  $q(\vec{r}, t)$ . Указанная энергия не является электромагнитной и не учитывается в балансе. Разогрев вещества в общем случае приводит к неравновесному процессу, причем разогретое вещество излучает во всем спектре, процесс является нестационарным, неравновесным и требует решения кинетического уравнения. Далее мы считаем процесс квазиравновесным и происходящим при постоянной температуре, т.е. интенсивности возбуждаемых полей считаем достаточно малыми, а теплоемкость вещества достаточно большой (бесконечно большой).

Уравнения баланса мощности получаются традиционно путем скалярного умножения каждого из уравнений (1) на вектор другого поля и сложения с использованием тождества  $\vec{a}(\nabla \times \vec{b}) - \vec{b}(\nabla \times \vec{a}) = -\nabla \cdot (\vec{a} \times \vec{b})$ . Здесь точкой обозначено скалярное произведение, а символом "×" – векторное. Указанную точку (если нет недоразумений) часто будем опускать. В результате имеем

$$\begin{aligned} \nabla \cdot \vec{S}(\vec{r}, t) + [\vec{E}(\vec{r}, t) \partial_t \vec{D}(\vec{r}, t) + \\ + \vec{H}(\vec{r}, t) \partial_t \vec{B}(\vec{r}, t)] = \\ = -[\vec{E}(\vec{r}, t) \vec{J}^e(\vec{r}, t) + \vec{H}(\vec{r}, t) \vec{J}^m(\vec{r}, t)]. \end{aligned} \quad (4)$$

В правой части (4) стоит плотность мощности, затрачиваемая источниками на создание поля. Это уравнение имеет форму

$$\begin{aligned} \nabla \cdot \vec{S}(\vec{r}, t) + \partial_t w(\vec{r}, t) = \\ = -[\vec{E}(\vec{r}, t) \vec{J}^e(\vec{r}, t) + \vec{H}(\vec{r}, t) \vec{J}^m(\vec{r}, t)], \end{aligned}$$

характерную для баланса [19]. Первый член есть плотность вытекающего потока мощности поля, а  $\partial_t w(\vec{r}, t)$  есть плотность накапливаемой мощности поля и вещества. Чтобы подсчитать затраченную работу, указанную величину надо проинтегрировать:

$$\begin{aligned} w(\vec{r}, t) = \partial_t^{-1} [\vec{E}(\vec{r}, t') \partial_t \vec{D}(\vec{r}, t') + \\ + \vec{H}(\vec{r}, t') \partial_t \vec{B}(\vec{r}, t')]. \end{aligned} \quad (5)$$

Именно эту величину, а не величину  $[\vec{E}(\vec{r}, t) \vec{D}(\vec{r}, t) + \vec{H}(\vec{r}, t) \vec{B}(\vec{r}, t)]/2$  (как обычно принято) следует ассоциировать с  $w(\vec{r}, t)$  [16]. Энергия (5), затраченная на поле и вещество, зависит от всей предыстории процесса, что для электродинамики сплошных сред естественно, тогда как для поля в вакууме этого не требуется [20]. Решая (3) совместно с (1), определим все поля во временном промежутке  $(0, t)$ . Это позволяет вычислить плотность диссипированной энергии  $q(\vec{r}, t)$  (выделенное тепло единицы объема). Заметим, что диссипация связана не только с проводимостью, но и с запаздыванием поляризации (т.е. отклика в виде индукций на воздействие полей). Окончательно для плотности энергии поле–вещество имеем  $e(\vec{r}, t) = w(\vec{r}, t) - q(\vec{r}, t)$ . Согласно концепции Н.Умова [21], скорость переноса энергии есть  $\vec{v}_e(\vec{r}, t) = \vec{S}(\vec{r}, t)/e(\vec{r}, t)$ . Она определена в каждой точке для каждого момента времени.

Уравнения Максвелла (3) теперь перепишем в виде

$$\begin{aligned} \nabla \times \vec{H}(\vec{r}, t) = \varepsilon_0 \partial_t \vec{E}(\vec{r}, t) + \vec{J}^e(\vec{r}, t) + \vec{J}_p^e(\vec{r}, t), \\ -\nabla \times \vec{E}(\vec{r}, t) = \mu_0 \partial_t \vec{H}(\vec{r}, t) + \vec{J}^m(\vec{r}, t) + \vec{J}_p^m(\vec{r}, t), \end{aligned}$$

а балансное уравнение так:

$$\begin{aligned} \nabla \cdot \vec{S}(\vec{r}, t) + \partial_t w_{(EM)}(\vec{r}, t) = \\ = -[\vec{E}(\vec{r}, t) (\vec{J}^e(\vec{r}, t) + \vec{J}_p^e(\vec{r}, t)) + \\ + \vec{H}(\vec{r}, t) (\vec{J}^m(\vec{r}, t) + \vec{J}_p^m(\vec{r}, t))], \\ w_{(EM)}(\vec{r}, t) = \partial_t^{-1} (\varepsilon_0 \vec{E}(\vec{r}, t') \partial_t \vec{E}(\vec{r}, t') + \\ + \mu_0 \vec{H}(\vec{r}, t') \partial_t \vec{H}(\vec{r}, t')) = \\ = [\varepsilon_0 \vec{E}^2(\vec{r}, t) + \mu_0 \vec{H}^2(\vec{r}, t)]/2. \end{aligned} \quad (7)$$



Поскольку  $\vec{S}$  есть плотность потока мощности поля и вещества, а уравнение (6) такое же, как для возбуждения поля в вакууме источниками в виде суммы сторонних токов и соответствующих токов поляризации, то смысл выражения (7) может быть в том, что это есть плотность собственной энергии поля. Величина  $w_{(EM)}$  в (6) определена с точностью до постоянной. Поскольку при  $t=0$  полная плотность энергии поля равна нулю, равна нулю и указанная постоянная. Величина вектора Пойтинга  $\vec{S}$  также определена с точностью до произвольного соленоидального вектора  $\vec{S}_0$ , т.е. такого, который удовлетворяет уравнению  $\nabla \cdot \vec{S}_0(\vec{r}, t) = 0$  [19]. Поток указанного вектора через любую замкнутую поверхность равен нулю, поэтому его влияние на общий поток энергии отсутствует. Однако можно показать, что нет никакой циркуляции энергии по замкнутым контурам. Поскольку при  $t=0$  поле отсутствовало, решать приведенное уравнение следует при условии  $\vec{S}_0(\vec{r}, 0) = 0$ , откуда имеем  $\vec{S}_0(\vec{r}, t) = 0$ . Для решения вопроса о скорости переноса собственной энергии поля, т.е. энергии, связанной только с фотонами (квазифотонами) [3], необходимо найти плотность потока энергии вещества. Если решена задача о движении частиц вещества в поле, можно определить среднюю его скорость  $\vec{v}(\vec{r}, t)$  в физически бесконечно малом объеме. Тогда, выбирая некий объем  $\Delta V$ , ограниченный поверхностью  $\Delta S$ , окружающей точку  $\vec{r}$ , в нерелятивистском пределе определим плотность потока так:

$$\begin{aligned} & \nabla \cdot \vec{S}_{(M)}(\vec{r}, t) = \\ & = \lim_{\Delta V \rightarrow 0} \frac{1}{2\Delta V} \oint_{\Delta S} \rho(\vec{r}', t) \vec{v}^2(\vec{r}', t) \vec{v}(\vec{r}') \cdot \vec{v}(\vec{r}', t) d^2 r'. \end{aligned}$$

Здесь предел следует понимать в смысле перехода к бесконечно малому объему. Отсюда имеем

$$\vec{S}_{(M)}(\vec{r}, t) = \left[ (1/2) \rho(\vec{r}, t) \vec{v}^2(\vec{r}, t) \right] \vec{v}(\vec{r}, t),$$

где  $\rho(\vec{r}, t)$  – плотность вещества,  $\vec{v}(\vec{r}, t)$  – скорость его движения. Соответственно

$\vec{S}(\vec{r}, t) = \vec{S}_{(M)}(\vec{r}, t) + \vec{S}_{(EM)}(\vec{r}, t)$ . В релятивистском случае следует использовать известную связь энергии с импульсом или со скоростью [22, 23]. Однако такой подход не очень конструктивен, поскольку он требует самосогласованно решать уравнения движения и уравнения возбуждения. На уровне макроскопического описания требуется решать уравнения возбуждения сторонними источниками и токами поляризации, а также и уравнения для векторов электрической и магнитной поляризации. По-видимому, используя только балансные соотношения, нельзя разделить баланс мощности или энергии на составляющие для поля и вещества. Для рассмотрения этого вопроса приведем уравнение (6) к виду

$$\begin{aligned} & \nabla \cdot \vec{S}_{(EM)}(\vec{r}, t) + \partial_t w_{(EM)}(\vec{r}, t) = \\ & = - \left[ \vec{E}(\vec{r}, t) \vec{J}^e(\vec{r}, t) + \vec{H}(\vec{r}, t) \vec{J}^m(\vec{r}, t) \right], \end{aligned} \quad (8)$$

$$\begin{aligned} & \nabla \cdot \vec{S}_{(EM)}(\vec{r}, t) = \nabla \cdot \vec{S}(\vec{r}, t) + \\ & + \vec{E}(\vec{r}, t) \vec{J}_p^e(\vec{r}, t) + \vec{H}(\vec{r}, t) \vec{J}_p^m(\vec{r}, t). \end{aligned} \quad (9)$$

Уравнение (8) такое же, как в вакууме. Вопрос стоит так: можно ли однозначно определить из (9) вектор  $\vec{S}_{(EM)}$ ? Согласно теореме Гельмгольца векторное поле  $\vec{S}_{(M)}$  представимо в виде суммы его потенциальной и соленоидальной частей:

$$\begin{aligned} & \vec{S}_{(M)}(\vec{r}, t) = \vec{S}(\vec{r}, t) - \vec{S}_{(EM)}(\vec{r}, t) = \\ & = \nabla \Phi(\vec{r}, t) + \nabla \times \vec{C}(\vec{r}, t). \end{aligned} \quad (10)$$

Отсюда имеем уравнение Пуассона для определения величины (10):

$$\begin{aligned} & \nabla^2 \Phi(\vec{r}, t) + \vec{E}(\vec{r}, t) \vec{J}_p^e(\vec{r}, t) + \\ & + \vec{H}(\vec{r}, t) \vec{J}_p^m(\vec{r}, t) = 0. \end{aligned} \quad (11)$$

Заметим, что эта величина плотности потока мощности определена с точностью до ротора вектора  $\vec{C}(\vec{r}, t)$ . Хотя указанный ротор есть вектор соленоидальный и поток не создает, так как его поток через любую замкнутую поверхность есть нуль, однако в общем случае  $\vec{C}(\vec{r}, t) \neq 0$ . Более того, может быть  $\vec{C}(\vec{r}, 0) \neq 0$ , т.е. в момент возникновения



поля в веществе могут циркулировать потоки. Следовательно, вектор (10) определен неоднозначно. Для того чтобы решить уравнение (11), необходимо предположить, что вещество находится в некотором ограниченном объеме  $V$  (что в нестационарном случае является требованием естественным). В этом случае, если известны  $\vec{E}(\vec{r}, t)$  и  $\vec{H}(\vec{r}, t)$ , решение может быть получено, например, методом функции Грина  $\Gamma(\vec{r}, \vec{r}') = (4\pi|\vec{r} - \vec{r}'|)^{-1}$  для уравнения Пуассона. Тогда собственная энергия поля переносится со скоростью  $\vec{v}_e^{(EM)}(\vec{r}, t) = \vec{S}_{(EM)}(\vec{r}, t)/w_{(EM)}(\vec{r}, t)$ . По-видимому, указанный подход нельзя распространить на бесконечную среду. Более того, скорость  $\vec{v}_e^{(EM)}(\vec{r}, t)$ , в отличие от скорости  $\vec{v}_e(\vec{r}, t)$ , не содержит в себе определенный физический смысл. Рассмотрим, например, плоскую монохроматическую волну в бесконечной среде. Амплитуды  $E_0$  электрического и  $H_0$  магнитного полей в ней связаны соотношением  $H_0 = Z_0 \rho_0 E_0$ , где  $Z_0 = \sqrt{\mu_0/\epsilon_0}$ , а  $\rho_0$  – действительный нормированный импеданс. В общем случае потерь в такой волне поля сдвинуты по фазе на угол  $\varphi$ . Для монохроматической волны могут выполняться условия

$$\langle \vec{E}(\vec{r}, t) \vec{J}_p^e(\vec{r}, t) \rangle \approx 0, \quad \langle \vec{H}(\vec{r}, t) \vec{J}_p^m(\vec{r}, t) \rangle \approx 0,$$

где скобки Дирака означают усреднение за период, т.е. поле в среднем не обменивается энергией с веществом. Эти уравнения выполняются строго в среде без дисперсии, т.е. для идеальных материальных соотношений  $\vec{D} = \epsilon_0 \epsilon \vec{E}$ ,  $\vec{B} = \mu_0 \mu \vec{H}$ . Тогда величиной (10) в среднем можно пренебречь, и вычислить величины  $\langle \vec{S} \rangle$  и  $\langle w_{(EM)} \rangle$ . Пусть волна движется вдоль направления единичного вектора  $\vec{z}_0$  оси  $z$ . Тогда мы должны были бы получить для скорости распространения чистой электромагнитной энергии выражение

$$\vec{v}_e^{(EM)} \approx \vec{z}_0 \frac{2c \cos(\varphi)}{\rho_0 + 1/\rho_0} \leq c. \quad (12)$$

В случае рассмотренных материальных уравнений имеем:

$$\varphi = 0, \quad \rho_0 = \sqrt{\mu/\epsilon}, \quad \vec{J}_p^e = \epsilon_0(\epsilon - 1)\partial_t \vec{E}, \\ \vec{J}_p^m = \mu_0(\mu - 1)\partial_t \vec{H}.$$

В случае  $\rho_0 = 1$ , т.е. при  $\epsilon = \mu$ , уравнение (12) дает скорость света, тогда как  $v_e = c/n = c/\epsilon$ . Если рассмотреть балансное соотношение (8), то видно, что второй член в правой части (9) естественным образом добавляется к  $w_{(EM)}$  и получается полная плотность энергии  $w = u = nw_{(EM)}$ , тогда как  $\vec{S}_{(EM)} = \vec{S}$ ,  $\vec{S}_{(M)} = 0$ .

Рассмотрим теперь баланс импульса. Первое уравнение в (3) умножим слева векторно на  $\vec{B}(\vec{r}, t)$ , а второе – на  $\vec{D}(\vec{r}, t)$  и вычтем одно из другого:

$$\begin{aligned} & [\vec{B}(\vec{r}, t) \times \nabla \times \vec{H}(\vec{r}, t) + \vec{D}(\vec{r}, t) \times \nabla \times \vec{E}(\vec{r}, t)] + \\ & + \partial_t (\vec{D}(\vec{r}, t) \times \vec{B}(\vec{r}, t)) = \quad (13) \\ & = -[\vec{J}^e(\vec{r}, t) \times \vec{B}(\vec{r}, t) + \vec{D}(\vec{r}, t) \times \vec{J}^m(\vec{r}, t)] = \\ & = -\vec{f}^L(\vec{r}, t). \end{aligned}$$

В правой части (13) стоит с обратным знаком сила Лоренца  $\vec{f}^L(\vec{r}, t)$ , действующая на сторонние токи, т.е. это сила, затрачиваемая сторонними источниками на создание импульса поля и вещества. Второй член слева в (13) есть производная плотности импульса поле–вещество по времени. Соответственно сама плотность импульса с точностью до постоянного вектора  $\vec{g}_0^M(\vec{r})$  есть

$$\vec{g}^M(\vec{r}, t) = \vec{D}(\vec{r}, t) \times \vec{B}(\vec{r}, t),$$

т.е. её следует взять в форме Минковского. Первый член в (13) представим в виде

$$\begin{aligned} \nabla \cdot \hat{\Sigma}(\vec{r}, t) &= \partial_v \hat{\Sigma}_v^v(\vec{r}, t) = \\ &= [\vec{B}(\vec{r}, t) \times \nabla \times \vec{H}(\vec{r}, t) + \\ &+ \vec{D}(\vec{r}, t) \times \nabla \times \vec{E}(\vec{r}, t)]. \end{aligned} \quad (14)$$



Здесь  $\nu = x, y, z$ . Величина  $\hat{\Sigma}(\vec{r}, t)$  есть тензор второго ранга в трехмерном пространстве. Поэтому дивергенция от него (свертка по одному индексу) есть стоящий в правой части (14) вектор. Указанный тензор также определен с точностью до произвольного тензора, для которого  $\nabla \cdot \hat{\Sigma}_0(\vec{r}, t) = 0$ . Поскольку при  $t = t_0 = 0$  поле отсутствовало, следует наложить начальные условия  $\vec{g}_0^M(\vec{r}) = 0$ ,  $\hat{\Sigma}_0(\vec{r}, 0) = 0$ . При указанных условиях в любой момент времени  $\vec{g}^M(\vec{r}, t)$  и  $\hat{\Sigma}(\vec{r}, t)$  определяются однозначно через решения задачи возбуждения, т.е. через поля  $\vec{E}(\vec{r}, t)$  и  $\vec{H}(\vec{r}, t)$  во все предыдущие моменты, при этом для определения  $\vec{g}^M(\vec{r}, t)$  следует вычислить интегралы (1), а для определения  $\hat{\Sigma}(\vec{r}, t)$  необходимо еще решить дифференциальное уравнение (14). Для решения (14) также можно воспользоваться теоремой Гельмгольца и решать уравнение Пуассона. Таким образом, уравнение баланса записывается так:

$$\partial_\nu \Sigma_\nu^v(\vec{r}, t) + \partial_t g_\nu^M = -f_\nu^L = -\partial_t \partial_t^{-1} f_\nu^L, \quad (15)$$

$$\nu' = x, y, z.$$

Здесь первый член есть поток компоненты  $\nu'$  полного импульса. Из этого уравнения следует, что скорость переноса компоненты  $\nu'$  импульса системы поле-вещество есть [19]

$$v_{i\nu'} = \partial_\nu \hat{\Sigma}_\nu^v(\vec{r}, t) / g_\nu^M, \quad (16)$$

а  $g_\nu^M$  – полный созданный импульс поля и вещества. Если рассмотреть все бесконечное пространство или объем, ограниченный поверхностью с радиусом  $r = ct$ , то сохраняется полный импульс поля, вещества и источника  $G_{\nu'} = \partial_r^{-1} g_\nu^M = \partial_r^{-1} (g_\nu^M + \partial_t^{-1} f_\nu^L) = 0$ , поскольку поток через поверхность равен нулю. Рассмотрим импульс, переданный веществу. Очевидно, указанная передача производится посредством токов поляризации

$$\vec{J}_p^e = \partial_t (\vec{D} - \epsilon_0 \vec{E}) = \sigma \vec{E} + N_p^e \partial_t \vec{p}^e,$$

$$\vec{J}_p^m = \partial_t (\vec{B} - \mu_0 \vec{H}) = N_p^m \partial_t \vec{p}^m.$$

Как уже было указано, ток проводимости  $\sigma \vec{E} = e \vec{v} N$ , где  $N$  – число носителей заряда в единице объема, учтен в токе поляризации. Соответственно  $N_p^e$  и  $N_p^m$  – числа электрических и магнитных диполей с моментами  $\vec{p}^e$  и  $\vec{p}^m$ . Переданный веществу удельный импульс есть

$$\vec{G}_{(M)} = \partial_t^{-1} \vec{g}_{(M)},$$

$$\vec{g}_{(M)}(\vec{r}, t) = \vec{J}_p^e(\vec{r}, t) \times \vec{B}(\vec{r}, t) + \vec{D}(\vec{r}, t) \times \vec{J}_p^m(\vec{r}, t). \quad (17)$$

Осталось найти плотность потока вещества. Для этого перепишем уравнения (3) в виде

$$\nabla \times \vec{H}(\vec{r}, t) = \epsilon_0 \partial_t \vec{E}(\vec{r}, t) + \vec{J}_p^e(\vec{r}, t) + \vec{J}^e(\vec{r}, t),$$

$$-\nabla \times \vec{E}(\vec{r}, t) = \mu_0 \partial_t \vec{H}(\vec{r}, t) + \vec{J}_p^m(\vec{r}, t) + \vec{J}^m(\vec{r}, t) \quad (18)$$

и запишем для этой формы уравнений Максвелла баланс импульса, умножив первое уравнение (18) слева векторно на  $\mu_0 \vec{H}(\vec{r}, t)$ , а второе – векторно на  $\epsilon_0 \vec{E}(\vec{r}, t)$  и вычтем одно из другого:

$$\nabla \cdot \hat{\Sigma}_{(EM)}(\vec{r}, t) + \partial_t \vec{g}^A(\vec{r}, t) = -\vec{f}^L(\vec{r}, t). \quad (19)$$

Здесь сила Лоренца

$$\vec{f}_{(M)}^L(\vec{r}, t) = \vec{J}_p^e(\vec{r}, t) \times \vec{B}(\vec{r}, t) + \vec{D}(\vec{r}, t) \times \vec{J}_p^m(\vec{r}, t),$$

действующая на вещество, перенесена в левую часть (19) и учтена в некой плотности потока  $\hat{\Sigma}_{(EM)}$ . Указанная плотность удовлетворяет дифференциальному уравнению

$$\nabla \cdot \hat{\Sigma}_{(EM)}(\vec{r}, t) = [\mu_0 \vec{H}(\vec{r}, t) \times \nabla \times \vec{H}(\vec{r}, t) + \epsilon_0 \vec{E}(\vec{r}, t) \times \nabla \times \vec{E}(\vec{r}, t)] + \vec{f}_{(M)}^L(\vec{r}, t) =$$

$$= \mu_0 [\nabla \vec{H}^2(\vec{r}, t) / 2 - (\vec{H}(\vec{r}, t) \cdot \nabla) \vec{H}(\vec{r}, t)] +$$

$$+ \epsilon_0 [\nabla \vec{E}^2(\vec{r}, t) / 2 - (\vec{E}(\vec{r}, t) \cdot \nabla) \vec{E}(\vec{r}, t)] + \vec{f}_{(M)}^L(\vec{r}, t). \quad (20)$$



Рассмотрим смысл баланса (19). Плотность Абрагама  $\vec{g}^A(\vec{r}, t)$  есть собственная плотность электромагнитного импульса поля. Она создается первичными (сторонними) источниками и вторичными источниками (токами поляризации вещества), определяющими  $\hat{\Sigma}_{(EM)}^A(\vec{r}, t)$ . При отсутствии источников ( $\vec{J}^L = 0$ ) уравнение (19) есть типичный закон сохранения. Следовательно, тензорная величина  $\hat{\Sigma}_{(EM)}^A(\vec{r}, t)$  определяет плотности потоков компонент собственного импульса поля. Она также определена с точностью до некоторого тензора  $\hat{\Sigma}_{(EM)}^{(0)}(\vec{r}, t)$ , дивергенция которого равна нулю и при начальном условии  $\hat{\Sigma}_{(EM)}^{(0)}(\vec{r}, 0) = 0$ . Поток импульса вещества определен тензором

$$\hat{\Sigma}_{(M)}^A(\vec{r}, t) = \hat{\Sigma}^A(\vec{r}, t) - \hat{\Sigma}_{(EM)}^A(\vec{r}, t).$$

Теперь можно определить скорости переноса собственного импульса поля и вещества соответственно:

$$\begin{aligned} v_{iv'}^{(em)} &= \partial_{v'} \hat{\Sigma}_{(EM)v'}^A(\vec{r}, t) / g_{v'}^A(\vec{r}, t), \\ v_{iv'}^{(M)} &= \partial_{v'} \hat{\Sigma}_{(M)v'}^A(\vec{r}, t) / g_{(M)v'}^A(\vec{r}, t). \end{aligned} \quad (21)$$

Для преобразования (20) было использовано векторное тождество

$$\begin{aligned} \nabla(\vec{a} \cdot \vec{b}) &= (\vec{a} \cdot \nabla)\vec{b} + (\vec{b} \cdot \nabla)\vec{a} + \\ &+ \vec{a} \times \nabla \times \vec{b} + \vec{b} \times \nabla \times \vec{a}, \end{aligned}$$

которое при  $\vec{a} = \vec{b}$  принимает вид

$$\nabla \vec{a}^2 = 2(\vec{a} \cdot \nabla)\vec{a} + 2\vec{a} \times \nabla \times \vec{a}.$$

Аналогично для преобразования введенных тензоров, например, тензора  $\hat{\Sigma}^A(\vec{r}, t)$ , можно использовать векторно-тензорное тождество

$$\begin{aligned} \vec{a} \times (\nabla \times \vec{b}) + \vec{b} \times (\nabla \times \vec{a}) &= \\ = \nabla \cdot [\hat{I}(\vec{a}\vec{b}) - \vec{a} \otimes \vec{b} - \vec{b} \otimes \vec{a}] + \\ + \vec{a}(\nabla \cdot \vec{b}) + \vec{b}(\nabla \cdot \vec{a}). \end{aligned}$$

Для одинаковых векторов оно приобретает форму

$$2\vec{a} \times (\nabla \times \vec{a}) = \nabla \cdot [\hat{I}\vec{a}^2 - 2\vec{a} \otimes \vec{a}] + 2\vec{a}(\nabla \cdot \vec{a}).$$

Для вакуума тензор  $\hat{\Sigma}^A(\vec{r}, t)$  равен максвелловскому тензору напряжений  $\hat{\sigma}_v^A$ , взятому с обратным знаком. Он также может быть преобразован с учетом того, что согласно (3)

$$\begin{aligned} \nabla \cdot \vec{D}(\vec{r}, t) &= -\partial_i^{-1}(\nabla \cdot \vec{J}^e(\vec{r}, t')) = \\ &= \partial_i^{-1}(\partial_i \rho^e(\vec{r}, t')) = \rho^e(\vec{r}, t), \\ \nabla \cdot \vec{B}(\vec{r}, t) &= -\partial_i^{-1}(\nabla \cdot \vec{J}^m(\vec{r}, t')) = \\ &= \partial_i^{-1}(\partial_i \rho^m(\vec{r}, t')) = \rho^m(\vec{r}, t), \end{aligned}$$

поскольку сторонние источники удовлетворяют уравнению непрерывности (закону сохранения зарядов):

$$\begin{aligned} \nabla \cdot \vec{J}^e(\vec{r}, t) + \partial_t \rho^e(\vec{r}, t) &= 0, \\ \nabla \cdot \vec{J}^m(\vec{r}, t) + \partial_t \rho^m(\vec{r}, t) &= 0. \end{aligned}$$

Так как сторонних магнитных зарядов нет, т.е.  $\rho^m(\vec{r}, t) = 0$ , плотность стороннего магнитного тока соленоидальна и может быть представлена как ротор от плотности некоторого электрического тока.

Итак, если плотность импульса определена по Минковскому, в недиспергирующей среде она есть плотность субстанции поле-вещество, а скорость ее переноса есть скорость фазовая. Её обобщение на диспергирующие среды приводит к скорости переноса (16), которая суть скорость переноса полного импульса поле-вещество. В этом случае для полного определения всех величин следует строго решать нестационарную задачу возбуждения, а все рассмотренные выше величины зависят от предыстории процесса, т.е. могут сложным образом зависеть от времени. Следует заметить, что полученные локальные (дифференциальные) балансные соотношения могут быть записаны в виде интегральных соотношений для некоторого объема  $V$ . Тогда интегралы от  $u$  и  $\vec{g}^M$  по указанному объему представляют собой соответственно полную энергию  $U$  и полный импульс  $\vec{G}^M$  указанного объема, которые сохраняются в смысле глобального закона сохранения. Может быть два случая. 1) В объеме есть источники поля. Тогда балансы полной энергии и импульса представляют



собой неоднородные балансные соотношения, в правые части которых входят величины, соответствующие производству энергии и импульса в объеме. Отрицательное производство энергии означает диссипацию. 2) В объеме не было источников до рассматриваемого момента времени  $t$ . В этом случае источники расположены вне объема, и в качестве момента  $t_0$  можно взять момент входа поля в объем. В этом случае энергия и импульс в объеме сохраняются в том смысле, что величина  $\partial_t(U + Q)$  в каждый момент времени равна вытекаемой из объема мощности, а изменение полного импульса объема равно вытекаемому из него импульсу. В случае стационарного (монокроматического) во времени поля или волны для ряда простых законов дисперсии можно получить конкретные виды указанных величин. В этом случае при предельном переходе от квазистационарного возбуждения к стационарному усредненные за период плотности «забывают свои начальные значения», т.е. перестают от них зависеть, а электромагнитный процесс выходит на стационарные значения своих величин. Для вакуума все полученные величины ТЭИ, плотности энергии и импульса совпадают с формой Абрагама. Заметим, что в нашем случае все величины определены однозначно. Так, решение дифференциальных уравнений  $\nabla \cdot \hat{\Sigma}_0(\vec{r}, t) = 0$  с нулевыми начальными условиями дает нулевые компоненты тензора  $\hat{\Sigma}_0(\vec{r}, t)$ .

## 2. Плоская монокроматическая волна в проводящем магнитодиэлектрике

Для более конкретного и простого анализа рассмотрим падение плоской монокроматической волны на слой магнитодиэлектрика толщины  $d$ . Пусть указанный слой расположен в области  $0 \leq z \leq d$  и имеет постоянные действительные спектральные проницаемости  $\varepsilon'$  и  $\mu$ , причем  $\varepsilon', \mu \geq 1$ . Это означает отсутствие у них частотной дисперсии, что справедливо в некоторой области частот  $0 < \omega \ll \omega_{\min}$ , где  $\omega_{\min}$  — некая минимальная частота из набора собст-

венных резонансных частот вещества, частоты нарушения нормального скин-эффекта (если таковой имеется), а также плазменной частоты свободных носителей заряда  $\omega_p$ .

Поскольку мы таковые рассматриваем, среда предполагается проводящей, т.е. имеющей комплексную спектральную диэлектрическую проницаемость с полюсом на нулевой частоте [17]

$$\varepsilon(\omega) = \varepsilon' - j\varepsilon'' = \varepsilon' - j\sigma/(\varepsilon_0\omega) \quad (22)$$

и магнитную проницаемость  $\mu = \text{const}$ , что мы далее и используем. Например, для воды  $\omega_{\min} \sim \omega_c = 1/\tau$  лежит в районе 100 ГГц,  $\omega_c$  — частота столкновений, а  $\tau$  определяет время релаксации в формуле Дебая. Для металлов  $\omega_{\min}$  может лежать в диапазонах от инфракрасного до ультрафиолетового [24, 25]. Далее мы также будем рассматривать плоскую квазимонокроматическую волну (длинный цуг или волновой пакет), подошедшую в момент  $t_0 = 0$  слева к границе слоя и возбужденную на бесконечности при  $t = -\infty$ . Формально математически такая плоская волна может быть возбуждена в обе стороны листом стороннего тока с плотностью  $\vec{J}_{inc}^e(x, y, z, t) = \vec{x}_0 \chi(t + \tau) I_x(t) \delta(z + l)$ , расположенного при  $z = -l$  и начавшего действовать в момент  $t = -\tau$  [26]. Здесь  $\chi$  — функция Хевисайда, а для квазимонокроматической волны удобно взять  $I(t) = \sin(\omega t)$ .

Далее мы рассмотрим дифракцию указанного цуга на пластине. Однако сначала рассмотрим плоскую квазимонокроматическую волну. Для нее плотность тока введем так:

$$\vec{J}_{inc}^e(x, y, z, t) = \vec{x}_0 \chi(t + \tau) \times I_x \delta(z + l) [1 - \exp(-\delta t)] \sin(\omega t), \quad (23)$$

и будем рассматривать поле при больших временах  $t \gg \tau + l/c$ . Как частный случай возбуждение с плотностью (23) будем также рассматривать в однородной проводящей бесконечной магнитодиэлектрической среде. Введенные плотности тока создают плоскую волну с компонентами  $E_x = E$  и  $H_y = H$  (да-



лее индексы  $x$  и  $y$  часто будем опускать). Указанная волна удовлетворяет уравнениям Максвелла в форме

$$\partial_z H = -\varepsilon_0 \varepsilon' \partial_t E - \sigma E, \quad \partial_z E = -\mu_0 \mu \partial_t H. \quad (24)$$

В вакууме  $\varepsilon' = \mu = 1, \sigma = 0$ .

Как известно, для плоской монохроматической волны с комплексной зависимостью  $\exp(j(\omega t - \beta(\omega)z) - \alpha(\omega)z)$ , движущейся вдоль  $z$  в диспергирующей среде, можно ввести две скорости, определяющие перенос присущих ей физических субстанций или ее характеристик: скорость движения энергии  $v_e$  и скорость движения электромагнитного импульса поля  $v_i$  [27–29] (в силу одномерности обозначение векторов опускаем). Кроме этого для закона дисперсии  $\beta(\omega)$  можно ввести еще две величины, определяющие движение математических (кинематических) характеристик волны: фазовую  $v_p(\omega) = \omega / \beta(\omega)$  и групповую  $v_g(\omega) = (\partial\beta(\omega) / \partial\omega)^{-1}$  скорости. Первая характеризует скорость движения фазы, а вторая – скорость движения фазовых возмущений или интерференционной картины (биений) двух бесконечно близких по частоте волн одинаковой амплитуды (как ее ввел Стокс). При этом положительное значение производной соответствует положительной дисперсии или прямой волне, а отрицательное – отрицательной дисперсии или обратной волне. Коэффициент  $n'(\omega) = c / v_p(\omega)$  определяет замедление и преломление волны на границе раздела, а коэффициент  $n''(\omega)$  – потери. При этом нормальной дисперсии соответствуют значения

$$\partial n'(\omega) / \partial\omega > 0 \text{ и } \partial v_p(\omega) / \partial\omega < 0,$$

а аномальной – соответственно [30]

$$\partial n'(\omega) / \partial\omega < 0 \text{ и } \partial v_p(\omega) / \partial\omega > 0.$$

Последние две скорости в общем случае не соответствуют движениям физических субстанций и являются лишь удобными математическими понятиями при описании волн [14–16, 31–41], хотя часто  $v_g(\omega)$  отождеств-

ляют с  $v_e(\omega)$ , что для диссипативных сред неверно [14–16, 33–35, 39–40]. Поскольку все реальные среды в той или иной мере имеют диссипацию,  $v_g(\omega) = v_e(\omega)$  выполняется лишь для ряда идеальных модельных случаев, например в идеальной бесстолкновительной плазме. Случай дисперсии в идеальном проводящих волноводах, включая и периодические волноводы, совсем иной: там при гармонической волне нет частотной спектральной группы волн, а закон дисперсии возникает из-за наличия пространственной спектральной группы парциальных волн, движущихся под углом  $\theta$  к оси волновода с фазовой скоростью, равной скорости света. Указанная группа в общем случае зависит от двух углов, один из которых определяет скорость переноса энергии вдоль оси, а другой может иметь непрерывный спектр [16]. При этом

$$v_p = c / \cos(\theta), \quad v_g = c \cos(\theta) \text{ и } v_p v_g = c^2.$$

Данный случай тривиален, соответствует электродинамическим структурам, а не средам, и здесь не рассматривается. В ряде работ, например, в [41], утверждается, что всегда  $v_g \leq c$ , хотя это не так [14–16, 39, 40]. Строго говоря, в задаче о распространении импульса (частотного волнового пакета или дуга) можно ввести бесконечное число величин, имеющих размерность скорости. Например, можно ввести такие скорости:  $v_n = (\omega^{n-1} \partial^n \beta(\omega) / \partial \omega^n)^{-1}$ . Для математического описания вводятся и соответствующие комплексные скорости, в которых величина  $\beta$  заменена комплексной постоянной распространения  $\gamma(\omega) = \beta(\omega) - j\alpha(\omega)$  [15, 37] и соответственно комплексный показатель преломления (коэффициент замедления)  $n(\omega) = n'(\omega) - jn''(\omega)$  [17]. Комплексная  $v_p$  для сигнала может иметь (и имеет) математический смысл, в отличие от комплексной  $v_g$ , которую можно ввести лишь для комплексных сигналов. В работе [42] введен 4-вектор групповой скорости, что имеет смысл в случае самосопряженного гамильтониана, когда





лагранжиан есть квадратичная по обобщенным координатам и импульсам функция. В литературе для консервативных (недиссипативных) систем известна теорема Леонтовича–Лайтхилла [27–29, 31, 43–45], согласно которой при указанном выше гамильтониане имеет место соотношение  $v_g = v_e$ . В нашем случае (и вообще для диссипативных сред) условия этой теоремы никогда не выполняются.

Традиционно  $v_g$  вводится путем разложения в спектральном интеграле фазовой постоянной  $\beta(\omega)$  в ряд Тейлора в окрестности некоторой частоты (например, несущей) с оставлением членов нулевого и первого порядков (первое приближение теории дисперсии) [36–38]. Иногда используют и обратное разложение  $\omega(\beta)$  [35]. Учет членов высших порядков как раз и приводит к появлению указанных выше скоростей, при этом вторая производная  $\partial^2 \beta / \partial \omega^2$  в первом приближении характеризует скорость расплывания импульса как целого [33–37]. В диссипативных средах можно ввести аналогичные комплексные скорости, зависящие также от постоянной затухания  $\alpha(\omega)$ . Указанные разложения являются асимптотическими [38, 46], т.е. не обязательно сходящимися. Сначала рассмотрим не распространение импульсов (цугов или волновых пакетов), а простейший одномерный случай плоской монохроматической волны. В такой волне нет частотной группы волн, поэтому и нет основания для введения групповой скорости (хотя формально для закона дисперсии  $\beta(\omega)$  ее ввести можно, что мы и делаем).

Часто необходимо рассматривать проводящие среды при достаточно низких частотах, когда дисперсией, не связанной с проводимостью, можно пренебречь. К ним, например, относятся полупроводники и металлы в радио и СВЧ диапазонах, морская вода в радиодиапазоне, ионосферная плазма (например, слой Хевисайда) при сверхнизких частотах. Указанный закон, когда дисперсия определяется только проводимостью  $\sigma$ , не зависящей от частоты (т.е. проводимостью на постоянном токе) имеет вид [35]

$$\beta(\omega) = (\omega/c) \sqrt{\epsilon' \mu / 2 \left[ 1 + \sqrt{1 + \sigma^2 / (\epsilon_0^2 \epsilon'^2 \omega^2)} \right]} = \omega n'(\omega) / c, \quad (25)$$

$$\alpha(\omega) = (\omega/c) \sqrt{\epsilon' \mu / 2 \left[ -1 + \sqrt{1 + \sigma^2 / (\epsilon_0^2 \epsilon'^2 \omega^2)} \right]} = \omega n''(\omega) / c. \quad (26)$$

Уже непосредственное применение формулы (22) демонстрирует, что групповая скорость может превышать величину  $c/\tilde{n}$ , где  $\tilde{n} = \sqrt{\epsilon' \mu}$  – коэффициент замедления или рефракции (показатель преломления) в среде в отсутствие проводимости. Более того, она может превышать и скорость света в вакууме  $c$ . Пусть  $\mu = 1$ . Обозначив  $\tilde{\omega} = \sigma / (\epsilon_0 \epsilon)$ , имеем:

$$\frac{d}{d\omega} \beta(\omega) = \frac{\beta(\omega)}{\omega} \times \left\{ 1 - \frac{(\tilde{\omega}/\omega)^2}{2 \left( 1 + \sqrt{1 + (\tilde{\omega}/\omega)^2} \right) \sqrt{1 + (\tilde{\omega}/\omega)^2}} \right\}_{\omega=\tilde{\omega}} = \frac{0.8535\dots}{v_p(\tilde{\omega})} = \frac{\tilde{n}/c}{1.0663\dots} \quad (27)$$

Указанное превышение наступает при  $\tilde{n} < 1.0663\dots$ , что может выполняться, например, в слабо ионизированном воздухе при условиях  $\omega_c \gg \omega_p$  и  $\omega_c \gg \tilde{\omega}$ . Введение в рассмотрение дисперсии диэлектрической проницаемости в ряде случаев позволяет сделать это превышение еще более существенным [16]. Сошлемся на две работы [39, 40], которые одними из первых отмечают факт превышения групповой скоростью скорости света в области аномальной дисперсии (см. также [15]) и в то же время показывают, что сигнал при этом распространяется со скоростью  $v < c$ . Таким образом, в проводящей среде  $v_g$  никак не может характеризовать скорость движения энергии. Наоборот, фазовая скорость в такой среде, обладающей аномальной положительной дисперсией, всегда меньше скорости света:

$$v_p(\omega) = c / n'(\omega) < c.$$



Целью данной работы также является доказательство для рассмотренного выше случая соотношений  $v_e = v_p$  и  $v_i = v_p/2$ . Заметим, что в работе [16] соотношение  $v_e = v_p$  для сред с аномальной положительной дисперсией, диэлектрическая проницаемость которых описывается формулой Дебая (т.е. для полярных диэлектриков с жесткими диполями), доказано двумя независимыми методами. Мы будем считать процесс распространения волны квазиравновесным, т.е. происходящим без разогрева при постоянной температуре.

### 3. Плотность электромагнитной энергии в монохроматической волне

Пусть плоская линейно поляризованная монохроматическая волна с компонентой электрического поля  $E_x$  распространяется в проводящей среде. Указанная волна создает плотность тока проводимости  $J_{x\sigma} = \sigma E_x$ , которая приводит к диссипации энергии волны. Эта диссипация имеет экспоненциальное распределение вдоль оси  $z$  вида  $\exp(-2\alpha(\omega)z)$  и создает неоднородный вдоль  $z$  нагрев бесконечного пространства. Неоднородный нагрев, в свою очередь, генерирует тепловое излучение в направлениях  $\pm z$ , имеющее все возможные спектральные компоненты. Таким образом, процесс изначально неравновесный. Его можно приближенно считать равновесным и одночастотным, если амплитуда волны на рассматриваемом участке мала (бесконечно мала) или же теплоемкость среды бесконечно велика.

В бесконечных диссипативных средах незатухающая гармоническая волна может распространяться только за счет энергии распределенных сторонних источников, которые компенсируют потери энергии волны на диссипированное тепло  $Q$  [17]. Мы предполагаем, что подобные источники находятся вне зоны рассмотрения волны (обычно на бесконечности). Для плоской волны энергия источника должна быть бесконечной даже в отсутствие затухания, что характеризует волну как удобную математическую абст-

ракцию (решение однородных уравнений Максвелла). Формально математически плоская волна возбуждается в обе стороны листом стороннего тока с плотностью (23), действующего бесконечно долго. Если же источники находятся на бесконечности, плоская волна есть предел сферической волны.

Диэлектрическая проницаемость (22) получается путем непосредственной подстановки экспоненциальных зависимостей для полей в уравнения Максвелла с учетом тока проводимости. Можно привести и другой вывод этой величины, а именно необходимо подсчитать среднюю за период поляризацию единицы объема и воспользоваться соотношением

$$D_x(\omega, t, z) = D(\omega, t, z) = \varepsilon_0 \varepsilon(\omega) E(\omega, t, z) = \varepsilon_0 E(\omega, t, z) + P_x(\omega, t, z). \quad (28)$$

Здесь  $E(\omega, t, z) = E_x(\omega, 0, 0) \exp(j\omega t - j\gamma(\omega)z)$ . Обозначим  $E_0 = E_x(\omega, 0, 0)$ . Соотношение (28) необходимо усреднить по периоду колебаний. В нашем случае  $\langle P_x(\omega, t, z) \rangle = (\kappa_1 + \kappa_2(\omega)) \langle E(\omega, t, z) \rangle$ . Для введенных восприимчивостей можно написать  $\kappa_1 = \varepsilon' - 1$ ,  $\kappa_2(\omega) = -j\sigma/(\varepsilon_0\omega)$ . Действительно, поляризация единицы объема в рассматриваемой среде создается поляризацией собственно вещества среды и движением свободных зарядов, рассеивающихся на атомах или молекулах вещества и друг на друге. Первая поляризация происходит без запаздывания (мгновенно) в силу нашего предположения, что характерные частоты вещества весьма высоки. Движение зарядов описывается уравнением  $\dot{x}(t) = \sigma(Ne)^{-1} \times E_x(\omega, t, z)$  и происходит так, что их потенциальная энергия равна нулю, а средняя кинетическая энергия имеет вид

$$\langle U_K \rangle = m\sigma^2 E_0^2 / (4Ne^2) \exp(-\alpha(\omega)z),$$

где  $\sigma = Ne^2/(m\omega_c)$ ,  $N$  – число заряженных частиц в единице объема. Этот результат также можно получить из диэлектрической восприимчивости плазмы



$$\kappa_p(\omega) = -\frac{\omega_p^2}{\omega(\omega - j\omega_c)} \quad (29)$$

в предположении, что  $\omega \ll \omega_p$  и  $\omega \ll \omega_c$ , т.е. что плазменная частота и частота столкновений весьма велики, при этом  $\sigma = \varepsilon_0 \omega_p^2 / \omega_c$  – проводимость при нулевой (бесконечно низкой) частоте. Иначе говоря, проводящую среду можно рассматривать как плазму при очень низких частотах. Усредненная плотность электрической части электромагнитной энергии и диэлектрическая проницаемость газа осцилляторов с собственной частотой  $\omega_0$  получены в [15] и имеют вид

$$\langle U_E(t, z) \rangle = \frac{\varepsilon_0 E_0^2}{4} \left\{ 1 + \frac{\omega_p^2(\omega^2 + \omega_0^2)}{(\omega^2 - \omega_0^2)^2 + \omega_c^2 \omega^2} \right\} \times \exp(-2\alpha(\omega)z), \quad (30)$$

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2(\omega^2 - \omega_0^2 + j\omega\omega_c)}{(\omega^2 - \omega_0^2)^2 + \omega_c^2 \omega^2}. \quad (31)$$

Для непроводящей среды без дисперсии следует положить  $\omega \ll \omega_0$  (или  $\omega = 0$ ), что дает

$$\varepsilon(\omega) = 1 + \omega_{p1}^2 / \omega_0^2 = 1 + \kappa_1 = \varepsilon' = \text{const}. \quad (32)$$

$$\begin{aligned} \langle U_E(t, z) \rangle &= (1/4)\varepsilon_0 E_0^2 (1 + \omega_{p1}^2 / \omega_0^2) \times \\ &\times \exp(-2\alpha(\omega)z) = \\ &= (1/4)\varepsilon_0 \varepsilon E_0^2 \exp(-2\alpha(\omega)z). \end{aligned} \quad (33)$$

Здесь индексом 1 отмечена плазменная частота, связанная с концентрацией диполей вещества. В такой среде энергия на частотах ниже  $\omega_{\min}$  распространяется с фазовой скоростью. Для проводящей среды с плазмой носителей следует добавить члены из формул (30), (31), в которых  $\omega_0 = 0$  (свободные заряды). В этом случае  $\omega_{p2} \gg \omega$ , поэтому

$$\langle U_E(t, z) \rangle = \frac{\varepsilon_0 E_0^2}{4} (1 + \omega_{p1}^2 / \omega_0^2 + \omega_{p2}^2 / \omega_c^2) \times \exp(-2\alpha z), \quad (34)$$

$$\begin{aligned} \varepsilon(\omega) &= 1 + \omega_{p1}^2 / \omega_0^2 - j\omega_{p2}^2 / (\omega_c \omega) = \\ &= 1 + \kappa_1 + \kappa_2. \end{aligned} \quad (35)$$

#### 4. Скорость движения энергии

Для любого волнового процесса скорость переноса некой субстанции определяется согласно [21] ее плотностью и вектором плотности ее потока в единицу времени. В нашем случае для энергии это вектор Пойтинга  $\vec{S}(z, t) = \vec{z}_0 S(z, t)$ , при этом:

$$\begin{aligned} \vec{v}_e(z, t) &= \vec{z}_0 v_e(z, t) = \vec{S}(z, t) / u(z, t) = \\ &= \vec{z}_0 E(z, t) H(z, t) / u(z, t), \end{aligned} \quad (36)$$

что следует из закона сохранения энергии (4) при общих предположениях. В (36) и далее использованы реальные физические поля. В случае гармонических полей в однородной среде, усредняя (36) по времени с учетом (30) и учитывая магнитную энергию, получим:

$$v_e(\omega) = c \frac{\sqrt{2}}{\sqrt{\varepsilon' \mu}} \frac{\sqrt{1 + \delta}}{1 + \zeta + \delta} < v_p(\omega), \quad (37)$$

где  $\delta = \sqrt{1 + \zeta^2}$ ,  $\zeta = \sigma / (\varepsilon_0 \varepsilon' \omega)$ . Если выполнены условия  $\omega \ll \omega_c$  и  $\sigma \gg \varepsilon_0 \varepsilon' \omega_c$ , то вторым слагаемым в знаменателе (37) можно пренебречь, и мы получим

$$v_e(\omega) \approx c \frac{\sqrt{2}}{\sqrt{\varepsilon' \mu} \sqrt{1 + \delta}} = v_p(\omega). \quad (38)$$

При указанных приближениях единицей под корнем также можно пренебречь, и тогда

$$\begin{aligned} v_e(\omega) &= v_p(\omega) = \frac{c}{\sqrt{\varepsilon' \mu}} \sqrt{\frac{2\varepsilon_0 \varepsilon' \omega}{\sigma}} = \\ &= \sqrt{\frac{2\omega}{\mu_0 \mu \sigma}} = v_g(\omega) / 2. \end{aligned} \quad (39)$$

Для  $\mu = 1$  в дисперсии (39) при небольших замедлениях возможно превышение групповой скоростью скорости света в пустоте. Именно из требования  $v_g > c$  имеем условие  $\sigma < 8\omega\varepsilon_0$ . Используя условие получения формулы (39), имеем  $\omega\varepsilon_0\varepsilon \ll \sigma < 8\omega\varepsilon_0$ . Для сред с  $\varepsilon \sim 1$  можно считать это неравенство выполненным при приближении  $\sigma$  к  $8\omega\varepsilon_0$  снизу. Однако при частотах порядка  $\sigma / (8\varepsilon_0)$  пренебречь полностью током сме-



щения по сравнению с током проводимости уже нельзя, поэтому следует рассматривать строгое соотношение (25).

Результат (38) можно получить другим способом – путем перехода от квазистационарного процесса к стационарному [16]. Пусть при  $z = 0$  в момент  $t = 0$  возник создающий поле источник поверхностного тока (23). При  $t < 0$  поле отсутствовало. В области  $|z| > 0$  он создает плоскую волну. Тогда плотность работы (энергии)  $w$ , затрачиваемой источником на создание поля (5), в нашем случае для наглядности запишем так:

$$\begin{aligned}
 w(z, t) = & \int_0^t \{ E(z, t') \partial_t D(z, t') + \\
 & + H(z, t') \partial_t B(z, t') \} dt' = \\
 = & \frac{\varepsilon_0 E^2(z, t) + \mu_0 H^2(z, t)}{2} + \\
 & + \int_0^t \{ \varepsilon_0 \hat{k}^e(0, z) E^2(z, t') + \\
 & + \mu_0 \hat{k}^m(0, \vec{r}) H^2(z, t') + \\
 & + \int_0^t [ \varepsilon_0 E(z, t') \partial_t \hat{k}^e(t' - t'', z) \vec{E}(t'', z) + \\
 & + \mu_0 \vec{H}(z, t') \partial_t \hat{k}^m(t' - t'', z) \vec{H}(z, t'') ] dt'' \} dt'.
 \end{aligned} \tag{40}$$

Здесь введены ядра интегральных операторов диэлектрической и магнитной восприимчивостей  $\hat{k}^e$  и  $\hat{k}^m$ . В силу однородности далее зависимость у них от  $z$  опускаем, а поля имеют экспоненциальную зависимость от  $z$ . Заметим, что в (40) в качестве нижнего предела можно взять  $-\infty$ . Для рассматриваемого закона дисперсии  $\hat{k}^m(t) = (\mu - 1)\delta(t)$ , а оператор диэлектрической восприимчивости имеет вид

$$\begin{aligned}
 \hat{k}^e(t) = & \frac{1}{2\pi} \int_{-\infty}^{\infty} \varepsilon(\omega) \exp(j\omega t) d\omega = \\
 = & (\varepsilon' - 1)\delta(t) + \chi(t)\sigma / \varepsilon_0,
 \end{aligned}$$

где  $\chi(t)$  – функция Хевисайда. Для определения плотности энергии поля из (40) необходимо вычесть плотность диссипированной энергии  $q$  (теплоты) [15–17]. Определяя эти

величины при больших временах  $t \gg 1/\delta$  и  $t \gg 1/\omega$ , получим указанные параметры для квазимонохроматического процесса. Усредняя по периоду колебаний  $2\pi/\omega$  и переходя к пределу  $t \rightarrow \infty$ , найдем энергию монохроматического процесса. Подставляя в интегралы типа (40) выражения всех входящих в них величин через спектральные интегралы, выделяя дельта-функции и интегрируя с ними по времени, а затем, вычисляя спектральный интеграл методом теории вычетов, придем к формуле (30). Заметим, что результат  $\langle U_E \rangle = \varepsilon_0 \varepsilon' |\vec{E}|^2 / 4$  для проводящей среды ( $\omega_0 = 0$ ) на малых частотах сразу следует из теоремы Умова–Пойтинга в комплексной форме [34,35]. Этот же результат следует из явного представления полей для уравнений (24) [35]:

$$\begin{aligned}
 E_x = E = E_0 \cos(\omega t - \beta z) \exp(-\alpha z), \\
 H_y = H = H_0 \cos(\omega t - \beta z - \varphi) \times \\
 \times \exp(-\alpha z).
 \end{aligned} \tag{41}$$

Здесь  $\varphi = \arctan(\alpha / \beta)$  – фазовый сдвиг, определяемый формулой (6.32) из [35], а отношения амплитуд в (18) дают действительный импеданс

$$Z = E_0 / H_0 = \omega \mu_0 \mu / \sqrt{\alpha^2 + \beta^2}.$$

При этом  $Z = Z_0 \sqrt{\mu / \sqrt{\varepsilon'^2 + \sigma^2} / (\omega^2 \varepsilon_0^2)}$  (формула (6.31) из [35]),  $Z_0 = \sqrt{\mu_0 / \varepsilon_0}$ . Отсюда следует

$$\begin{aligned}
 v_e = \frac{\langle S \rangle}{\langle U \rangle} = & 2 \frac{E_0^2 \cos(\varphi) / Z}{\varepsilon_0 \varepsilon' E_0^2 + \mu_0 \mu E_0^2 / Z^2} = \\
 = & \frac{2}{\omega \mu_0 \mu \varepsilon_0 \varepsilon' + \varepsilon_0 \varepsilon' \sqrt{1 + \zeta^2}} = \\
 = & \frac{2}{\omega^2 \varepsilon_0 \varepsilon' \mu_0 \mu (1 + \delta)} = \frac{\omega}{\beta} = v.
 \end{aligned}$$

### 5. Скорость переноса импульса поля

Рассмотрим вопрос о скорости переноса электромагнитного импульса поля. В вакууме плотность импульса определяется од-



нозначно вектором Абрагама  $\vec{g}^A(t, z) = \vec{z}_0 S(t, z) / c^2$ . В сплошной среде традиционно имеются две формы: Абрагама  $\vec{g}^A$  и Минковского

$$\vec{g}^M(t, z) = \vec{z}_0 \tilde{n}^2 S(t, z) / c^2 = \tilde{n}^2 \vec{g}^A(t, z),$$

причем до сих пор вопрос в пользу выбора одного из них не решен [3–13, 41], что даже отражено в физической энциклопедии. В отсутствие дисперсии оба определения приводят к одинаковым законам сохранения для соответствующих тензоров энергии–импульса (ТЭИ) поля в форме Минковского и Абрагама, которые эквивалентны [3]. Считается, что выбор вида для плотности импульса невозможен без решения уравнений движения вещества в поле и определения ТЭИ среды. Однако без определения плотности импульса и его потока, а также хотя бы одной из этих величин невозможно определить и скорость переноса импульса. Подобные неоднозначности вообще характерны для электродинамики сплошных сред. Так, введение электродинамических потенциалов в среде также неоднозначно [47], что связано не только с калибровочными преобразованиями. В нашем случае плоской волны  $\vec{g}^M(t, z)$  и  $\vec{g}^A(t, z)$  имеют вид

$$\begin{aligned} \vec{g}^A(t, z) &= \vec{z}_0 g^A(t, z) = \\ &= \frac{\vec{z}_0 E_0^2}{c^2 Z} \cos(\omega t - \beta(\omega)z) \times \end{aligned} \quad (42)$$

$$\times \cos(\omega t - \beta(\omega)z - \varphi) \exp(-2\alpha(\omega)z),$$

$$\vec{g}^M(t, z) = \vec{z}_0 g^M(t, z) =$$

$$= \frac{\vec{z}_0 \tilde{n}^2 E_0^2}{c^2 Z} \cos(\omega t - \beta(\omega)z) \times \quad (43)$$

$$\times \cos(\omega t - \beta(\omega)z - \varphi) \exp(-2\alpha(\omega)z).$$

Мы введем еще одно определение, которое ближе к определению Минковского:  $\vec{g} = \vec{z}_0 \tilde{g}^M$ , где  $\tilde{g}^M = S / v_p^2 = n'^2(\omega) g^A$ . При  $\sigma = 0$  имеем  $g^M = \tilde{g}^M$ ,  $\alpha = 0$ ,  $\beta = \omega \sqrt{\epsilon' \mu} / c$ , при этом импульс в среднем веществе не передается (за исключением случаев отраже-

ния от границ разделов), поскольку сила Лоренца и ток электрической поляризации  $\vec{J}_p^e = \epsilon_0(\epsilon' - 1) \partial \vec{E} / \partial t$  сдвинуты по фазе на  $\pi/2$ , и импульс движется с фазовой скоростью  $v_i = v_p = c / \tilde{n}$  [29]. Это же относится к силе Лоренца  $\epsilon_0 \epsilon' \vec{E} \times \vec{J}_p^m$  для магнитного тока поляризации  $\vec{J}_p^m = \mu_0(\mu - 1) \partial \vec{H} / \partial t$ .

В диссипативных средах за счет поглощения фотонов происходит передача импульса веществу, поток импульса направлен по  $z$ , поэтому уравнение баланса для вектора плотности импульса можно записать в виде [19, 29]

$$-\partial_t [g(t, z) + g_\sigma(t, z)] = \partial_z [v_i(t, z)g(t, z)] = \quad (44)$$

$$= v_i(t, z) \partial_z g(t, z),$$

где введен переданный веществу импульс  $g_\sigma(t, z)$ . Здесь предположено, что в силу однородности среды  $v_i$  не зависит от  $z$ . Для любой из трех форм имеем функциональную зависимость  $g(t, z) = G(\omega t - \beta z) \exp(-2\alpha z)$ . Скорость изменения импульса, передаваемого слою вещества толщины  $dz$  с единичной площадью, есть разность потоков  $g$  через сечения  $z$  и  $z + dz$ , поэтому

$$\partial_t g_\sigma(t, z) = 2\alpha v_i(t, z)g(t, z). \quad (45)$$

Величина (45) равна давлению поля на слой единичной толщины. Для получения давления на конечный слой (45) следует интегрировать в пределах его координат. В бесконечном слое весь импульс поля передается веществу. Подставляя (45) в (44), получим, что скорость переноса импульса постоянна и равна половине фазовой скорости:

$$v_i(t, z) = -\partial_t g(t, z) / [\partial_z g(t, z) + 2\alpha(\omega)g(t, z)] = \omega / 2\beta(\omega) = v_p(\omega) / 2. \quad (46)$$

Здесь форма записи импульса не конкретизирована. Это же соотношение можно получить из следующего рассуждения. Уравнения Максвелла записываются в форме (24). Отсюда следует уравнение баланса

$$\begin{aligned} \partial_z [\epsilon_0 \epsilon' E^2 / 2 + \mu_0 \mu H^2 / 2] = \\ = -\partial_t g^M - \mu_0 \mu \sigma E H. \end{aligned} \quad (47)$$



Поскольку  $\partial S / \partial t = -v_p (\partial S / \partial z + 2\alpha S)$ , то единственный вид  $g = \eta S / c^2$ , компенсирующий действующую на заряды силу Лоренца  $\mu_0 \mu \sigma S$ , получается при форме  $\tilde{g}^M$ . Действительно, требуем выполнения уравнения баланса  $2\alpha \eta S v_p / c^2 = \mu_0 \mu \sigma S$ . Указанный баланс записан для зарядов среды, откуда  $\eta = c^2 \omega \mu_0 \mu \sigma \beta / (2\alpha \omega^2) = c^2 / v_p^2$ . Поэтому в правой части (47) делаем замену  $\partial g^M / \partial t = \partial \tilde{g}^M / \partial t + \tilde{f}$ , где  $\tilde{f} = c^{-2} (\tilde{n}^2 - n'^2) \partial S / \partial t$ . Средняя плотность энергии в (47) имеет вид  $\langle u \rangle = \langle S \rangle / v_p = E_0^2 \cos(\varphi) / (2Zv_p) = E_0^2 / (2v_p^2 \mu_0 \mu)$ . Аналогично для третьей формы импульса имеем  $\langle \tilde{g}^M \rangle = n'^2 \langle S \rangle / c^2 = n'^2 E_0^2 / (2v_p \mu_0 \mu)$ . Поэтому скорость переноса импульса  $\tilde{g}^M$  есть

$$\tilde{v}_i^M = \langle u \rangle / \langle \tilde{g}^M \rangle = c^2 / (n'^2 v_p) = v_p(\omega) / 2. \quad (48)$$

В этом случае в уравнении баланса появляется дополнительный член  $\tilde{f}$ . Он подобен силе Абрагама и действует на среду. Однако появление указанной силы (как и силы Абрагама) не должно приводить к каким-то возражениям по невыполнению уравнений баланса, поскольку  $\langle \partial S / \partial t \rangle = 0$  для любого фазового сдвига  $\varphi$ . Поэтому как указанная сила, так и сила Абрагама импульс веществу в среднем не передают. В непроводящей среде  $\tilde{f} = 0$ ,  $\tilde{g}^M = g^M$ , а энергия и импульс распространяются с фазовой скоростью  $v_p = \omega / \beta = c / \tilde{n}$ , не зависящей от частоты. Перенос дивергентную часть  $\tilde{g}^M$  в левую часть равенства (47), получаем другое уравнение баланса  $\langle u - S / v_p \rangle = 0$ , которое есть тождество. Оставшийся член в его правой части  $\tilde{f}$  есть сила, действующая на среду, но не передающая в среднем за период ей импульс. Заметим, что использование тради-

ционного вектора плотности импульса по Минковскому не приводит к какому-либо физически ясному выражению для скорости его переноса в проводящей среде. Однако если есть локализованный сторонний электрический ток, а  $\sigma = 0$ , то введение  $\tilde{g}^M$  более удобно и приводит к переносу полного импульса со скоростью  $v_i = v_p = v_g = v_e = c / \tilde{n}$ , что тривиально, поскольку при этом нет дисперсии. Собственно так традиционно и вводится ТЭИ в среде [3–5].

В большом числе работ принято определять правильную плотность импульса по Абрагаму  $g^A(t, z) = S(t, z) / c^2$  [3–5]. При этом дополнительно возникает объемная сила Абрагама  $f^A = c^{-2} (\epsilon' \mu - 1) \partial S / \partial t$ . Такое определение считается более строгим, хотя для электродинамики сплошных сред определение Минковского часто более удобно [3,5] и соответствует экспериментам по давлению света. Уравнение баланса (47) можно переписать так:

$$\begin{aligned} \partial_z [\epsilon_0 \epsilon' E^2 / 2 + \mu_0 \mu H^2 / 2] = \\ = -\partial_i g^A - f^A - \mu_0 \mu \sigma S. \end{aligned} \quad (49)$$

Теперь изменение плотности импульса в секунду, воздействующее на вещество, есть  $(2\alpha \omega \epsilon' \mu / \beta) S / c^2$ . Это есть результат вклада первого и второго членов в правой части (49). Но оно не компенсирует последнее слагаемое, поскольку  $\tilde{n} \neq n'$ . Очевидно, использование первого или второго членов в правой части (49) порознь тоже не приводит к какой-либо подобной компенсации. Используя в качестве плотности импульса величину  $g^A$ , также не удастся получить какое-либо разумное определение скорости его движения, которая бы не превышала  $c$  в случае среды с дисперсией. Исключение составляет случай непроводящей недисперсивной среды с  $\tilde{n} = n'$ . Поскольку  $c^{-2} (\epsilon' \mu - 1) \partial_i S = -v_p c^{-2} (\epsilon' \mu - 1) \partial_z S$ , то, внося дивергентную часть  $f^A$  в плотность потока, имеем:



$$\begin{aligned}
 v_i^A &= \frac{\langle u - v_p (\varepsilon' \mu - 1) S / c^2 \rangle}{\langle g^A \rangle} = \\
 &= c^2 \langle [S] / v_p - v_p (\varepsilon' \mu - 1) [S] / c^2 \rangle / \langle S \rangle = \\
 &= c \tilde{n} - v_p (\tilde{n}^2 - 1) = v_p = c / \tilde{n}.
 \end{aligned}$$

Таким образом, и в этом случае импульс поля переносится с фазовой скоростью, которая меньше скорости света и совпадает со скоростью энергии в недиспергирующей среде.

Рассмотрим передаваемый среде в единицу времени (секунду) импульс. Для этого запишем уравнение баланса в форме

$$\begin{aligned}
 \partial_z [\varepsilon_0 E^2 / 2 + \mu_0 H^2 / 2] = \\
 = -\partial_t g^A - \mu_0 \mu \sigma S + \mu_0 (\mu - 1) \sigma S - \quad (50) \\
 - c^2 (\varepsilon' - 1) H \partial_t E - c^2 (\mu - 1) E \partial_t H.
 \end{aligned}$$

Оно также получено из уравнений Максвелла, когда среда учитывается в виде токов поляризации: электрического  $J_{px}^e = \varepsilon_0 (\varepsilon' - 1) \partial_t E$  и магнитного  $J_{py}^m = \mu_0 (\mu - 1) \partial_t H$ . Такой подход более присущ микроскопической электродинамике [7, 8], при этом описание становится весьма сложным. Известно, что сила Абрагама не полностью характеризует воздействие на среду [5]. Три последних члена в (50) правильно отражают такое воздействие. Именно  $f_{\sigma}^L = \mu_0 \mu \sigma S = \sigma E B$  есть сила Лоренца, действующая на ток проводимости,  $f_{pe}^L = c^{-2} (\varepsilon' - 1) H \partial_t E = J_{px}^e B$  есть сила Лоренца, действующая на ток электрической поляризации среды, а  $f_{pm}^L = c^{-2} (\mu - 1) E \partial_t H = D_x J_{py}^m$  — сила Лоренца, действующая на ток магнитной поляризации среды (здесь  $B = \mu_0 \mu H$ ). Для первой силы мы получили усредненное значение  $\langle f_{\sigma}^L \rangle = \mu_0 \mu \sigma \langle S \rangle = \sigma E_0^2 \beta / 2 \omega = n' \sigma E_0^2 / (2c)$ . Для второй и третьей сил соответственно имеем  $\langle f_{pe}^L \rangle = -(\varepsilon' - 1) E_0^2 \sigma / (4n'c)$  и  $\langle f_{pm}^L \rangle = (\mu - 1) \times E_0^2 \sigma / (4n'c)$ . Поскольку обычно  $n' \gg 1$ ,

первый член играет решающую роль в передаче импульса подвижным зарядам. Они, рассеиваясь на молекулах и атомах вещества, передают импульс ему. В единицу времени через единичную поверхность переносится энергия  $S = v_p u$ . Теряемая в единице объема мощность  $\sigma E_0^2 \cos^2(\omega t - \beta z)$ , усредненная за период, есть  $\sigma E_0^2 / 2$ . С другой стороны, эту же величину можно определить так:  $\langle \partial u_{\sigma} / \partial t \rangle = v_p^{-1} \langle \partial S_{\sigma} / \partial t \rangle$ , что равно  $2\alpha \langle S \rangle = \alpha E_0^2 \cos(\varphi) / Z = \sigma E_0^2 / 2$ . Потерянной энергии соответствует переданный импульс  $\langle \partial g_{\sigma} / \partial t \rangle$ . Для трех введенных определений плотности импульса соответственно имеем:

$$\begin{aligned}
 \langle \partial_t g_{\sigma}^M \rangle &= 2\alpha \tilde{n}^2 v_p E_0^2 \cos(\varphi) / (c^2 Z) = \\
 &= \sigma \tilde{n}^2 E_0^2 v_p / (2c^2), \\
 \langle \partial_t g_{\sigma}^A \rangle &= 2\alpha v_p E_0^2 \cos(\varphi) / (c^2 Z) = \\
 &= \sigma \tilde{n}^2 E_0^2 v_p / (2c^2), \\
 \langle \partial_t \tilde{g}_{\sigma}^M \rangle &= 2\alpha n'^2 v_p E_0^2 \cos(\varphi) / (c^2 Z) = \\
 &= \sigma E_0^2 / (2v_p).
 \end{aligned}$$

Поскольку свободные заряды удовлетворяют уравнению движения

$$Nex(t) = \sigma E_0 \cos(\omega t - \beta(\omega)z) \exp(-\alpha(\omega)z),$$

может показаться, что у вещества есть колеблющаяся  $x$ -компонента импульса. Однако это не так. Коль скоро мы предполагаем вещество электрически нейтральным, всегда имеются заряды, смещающиеся в противоположную сторону, которым и передается противоположный импульс (например, в металлах это кристаллическая решетка). Может возникнуть вопрос: откуда взялся у поля, возбуждаемого диполем или системой диполей (плоскостью диполей в нашем случае) импульс, коль скоро источники его не имели? Здесь ответ тривиален: источник возбуждает две волны в сторону положительных и отрицательных  $z$ , импульсы которых равны и противоположны.



Следует заметить, что в веществе элементарные кванты импульса поля переносятся фотонами между актами их рассеяния на частицах вещества со скоростью света  $c$ . Перенос импульса с фазовой скоростью есть коллективный результат указанных элементарных актов с учетом соответствующих фазовых задержек и интерференций. Формально учет влияния вещества может быть осуществлен введением токов поляризации. Для недиссипативных сред (т.е. в отсутствие дисперсии  $\varepsilon'$  и  $\mu$ ) сдвиг между ними и полями есть  $\pi/2$ , при этом импульс среде посредством этих токов не передается, а передается только с помощью тока проводимости. Уравнение баланса (50) (и аналогичное для баланса мощности) с введением плотности  $u_0 = (\varepsilon_0 E^2 + \mu_0 H^2)/2$  также не удобно для определения явлений переноса в монохроматической волне в среде.

Балансные соотношения с введением токов поляризации и сторонних токов, создающих поле, весьма продуктивны при нестационарном возбуждении [16]. В этом случае накопленная в некотором объеме энергия и импульс зависят от предыстории процесса создания поля. Например, в плазме следует учитывать энергию и импульс системы поле-вещество. Для плазмы это означает учет кинетической энергии колебаний заряженных частиц и передаваемый частицам импульс. Решая уравнения движения системы поле-вещество, в принципе можно определить ТЭИ поля и вещества в любой момент и соответственно мгновенные величины  $\vec{v}_e(\vec{r}, t)$  и  $\vec{v}_i(\vec{r}, t)$ . В случае макроскопической электродинамики сплошной среды усреднение по физически бесконечно малому объему (гомогенизация) приводит к материальным уравнениям, являющимся аналогом уравнений движения. Например, одномерные нестационарные уравнения (24) в однородной среде запишем с учетом только временной (частотной) дисперсии:

$$\begin{aligned} -\partial H / \partial z &= \partial D / \partial t + J^e, \\ -\partial E / \partial z &= \partial B / \partial t + J^m. \end{aligned} \quad (51)$$

Все величины в (51) есть функции времени и одной координаты  $z$ , что упрощает рассмот-

рение. Возьмем материальные уравнения, определяющие дисперсию, в форме

$$\begin{aligned} D(z, t) &= \varepsilon_0 \int_0^t \varepsilon(z, t-t') E(z, t') dt', \\ B(z, t) &= \mu_0 \int_0^t \mu(z, t-t') H(z, t') dt'. \end{aligned} \quad (52)$$

Не уменьшая общности, будем считать, что поле отсутствовало при  $t=0$ . Если источники расположены в плоскости  $z=0$ , то в момент  $t$  поле будет расположено в пределах  $|z| \leq ct$ . Наличие проводимости соответствует полюсу в нуле у спектральной комплексной диэлектрической проницаемости

$$\begin{aligned} \varepsilon(z, \omega) &= \int_{-\infty}^{\infty} \varepsilon(z, t) \exp(-j\omega t) dt = \\ &= \int_0^t \varepsilon(z, t') \exp(-j\omega t') dt'. \end{aligned}$$

Здесь учтено, что  $\varepsilon(z, t) = 0$  при  $t < 0$ . Наиболее удобно использовать модель плазмы. Для волны в плазме можно ввести затухание Ландау и исключить указанный полюс. Записывая уравнение баланса для импульса, получим такую форму:

$$\begin{aligned} \partial_z u_0(z, t) &= -[\mu_0 H \partial_t D(z, t) + \varepsilon_0 E \partial_t B(z, t)] - \\ &\quad - \mu_0 H J^e - \varepsilon_0 E J^m. \end{aligned} \quad (53)$$

Последние два члена и создают импульс поля и вещества. Однако они соответствуют силе Лоренца в вакууме, а не в веществе. Подразумевая, что в квадратной скобке должна стоять величина типа  $\partial_t g(z, t)$ , получаем после интегрирования соотношение

$$\begin{aligned} g(z, t) &= \partial_t^{-1} (\varepsilon_0 E(z, t') \partial_t B(z, t') + \\ &\quad + \mu_0 H(z, t') \partial_t D(z, t')) = \\ &= \varepsilon_0 \varepsilon(z, 0) E(z, t) + \mu_0 \mu(z, 0) H(z, t) + \\ &\quad + \partial_t^{-1} (\varepsilon_0 \partial_t \varepsilon(z, t-t') E(z, t') + \\ &\quad + \mu_0 \partial_t \mu(z, t-t') H(z, t')) = \\ &= \mu_0 H(z, t) D(z, t) + \varepsilon_0 E(z, t) B(z, t) - \\ &\quad - \partial_t^{-1} (\mu_0 \partial_t H(z, t') D(z, t') + \\ &\quad + \varepsilon_0 \partial_t E(z, t') B(z, t')). \end{aligned} \quad (54)$$





Из него следует, что эта  $z$ -компонента (54) в момент  $t$  зависит не только от значений полей в данный момент, но и от всех предыдущих их значений. Волны в однородной плазме  $\mu(z, t) = \delta(t)$ ,  $\varepsilon(z, t) = \delta(t) + (\omega_p^2 / \omega_c) \times (1 - \exp(-\omega_c t)) \exp(-\omega_L t)$ , где последняя экспонента определяет затухание Ландау, причем в окончательных результатах можно считать  $\omega_L \rightarrow 0$ . Тогда  $\varepsilon(\omega) = 1 + \omega_p^2 / (\omega^2 - j\omega\omega_c)$ . Задавая плотность электрического тока в виде  $J^e = I[1 - \exp(-t/\tau)] \sin(\omega t)$ , можно решить уравнения (51) и определить плотность (54) при больших значениях  $t$ , когда процесс становится квазистационарным. Однако соотношение (54) не соответствует импульсу в сплошной среде. Для вакуума оно становится тривиальным:  $g = g^A = S/c^2$ . Сплошной среде более соответствует плотность  $g = g^M = DB$  и уравнение баланса импульса, имеющее для (51) вид

$$\partial_z \Sigma + \partial_t g^M = -B_y J_x^e - D_x J_y^m. \quad (55)$$

Здесь  $\partial_z \Sigma = B \partial_z H + D \partial_z E$ ,  $g^M = DB$ ,  $\Sigma$  – плотность потока импульса в направлении  $z$ , а в правую часть (55) входят реальные силы Лоренца с учетом коллективного влияния движения всех зарядов сплошной среды. По-видимому, соотношение (55) более приемлемо для сплошной среды, поскольку при отсутствии дисперсии, т.е. при материальных соотношениях  $D = \varepsilon_0 \varepsilon E$ ,  $B = \mu_0 \mu H$ , имеем  $g^M = \varepsilon \mu S / c^2$ ,  $\Sigma = u = (\varepsilon_0 \varepsilon E^2 + \mu_0 \mu H^2) / 2$ ,  $v_i^M = c / \sqrt{\varepsilon \mu} = v_p$ , т.е. мы получили обобщение плотности импульса в форме Минковского. Трудность в нестационарном случае состоит в том, что вычисление  $g^M = DB$ , как и  $\Sigma$ , требует знания всей предыстории электромагнитного процесса, при этом  $\Sigma$  еще надо находить из решения дифференциального уравнения. В одномерном случае дифракции на пластине для этого достаточно выполнить интегрирование по  $z$ . Введение же силы Абрагама в (55) как довесок к плотности импульса по Абрагаму (чтобы факти-

чески получить импульс по Минковскому) только запутывает рассмотрение. Приведем соответствующее (51) аналогичное уравнение баланса мощности  $p = \partial_t w = \partial_t (u + q)$ :  $\partial_t w + \partial_z S = -E J_x^e - H J_y^m$ . Здесь  $q(z, t)$  – диссипированная работа источников поля (тепло), которую всегда можно вычислить. Для простейшего случая проводящей среды  $q(z, t) = \sigma \partial_t^{-1} (E^2(z, t'))$ . Отсюда следует выражение для плотности энергии поле-вещество:

$$u(z, t) = \int_0^t (E(z, t') \partial_t D(z, t') + H(z, t') \partial_t B(z, t')) dt' - q(z, t). \quad (56)$$

#### Заключение

Получены представления плотностей энергии и импульса в плоской монохроматической волне для дисперсии, определяемой только током проводимости. Показано, что в этом случае энергия переносится с фазовой скоростью, которая меньше скорости света  $c$  в вакууме, тогда как групповая скорость может превышать скорость света. Импульс при малой  $\sigma$  переносится с фазовой скоростью, а при большой – с половинной фазовой скоростью. Рассмотрено несколько форм представления плотности импульса и показано, что форма Минковского предпочтительнее, чем форма Абрагама. Результат для скорости переноса энергии получен несколькими способами, в частности использован закон дисперсии для газа осцилляторов. Выводы обобщаются и при наличии нескольких резонансных частот осцилляторов, а также и для наличия внутреннего поля. Существенным здесь является подсчет поляризуемости с использованием уравнения движения частиц, которое имеет первый порядок. В этом случае потенциальная энергия в веществе не накапливается, а его собственные колебания отсутствуют. Приведенные результаты могут быть обобщены для проводящих полярных диэлектриков, например для воды, содержащей ионы проводимости. В идеально дистиллированной воде замедление изменяется от 9 примерно до 1, а скорость энергии сов-



падает с фазовой скоростью. В содержащей ионы проводящей воде в области сверхнизких частот коэффициенты замедления и потерь стремятся к бесконечности. В морской воде ( $\sigma = 4$  С/м) на частоте около 900 МГц ток смещения равен току проводимости, а на существенно более низких частотах вода ведет себя подобно металлу. Знание скорости движения энергии и соответственно импульсов важно для передачи сообщений, например, при связи с подводными лодками, которая осуществляется на сверхнизких частотах, при этом сигнал передается со скоростью, примерно равной фазовой, а групповая скорость в два раза больше. Например, импульс с несущей частотой 1 кГц и длительностью 2 мс (два периода) достигает подводной лодки на глубине 100 м от поверхности за 2 мс и детектируется по максимуму огибающей (без учета времени срабатывания детектора), тогда как сильно размытый фронт уширенного в несколько раз импульса приходит через 1 мс и не может быть протектирован [38].

Вопрос о форме ТЭИ в электродинамике сплошных сред и о соответствующих плотностях, известный как Абрагама–Минковского контroversия, до сих пор остается дискуссионным (см. дополнительно [48–53]). В различных модельных примерах как за определенную форму ТЭИ, так и против неё, часто содержатся неточности, разбирать которые в данной работе нет возможности (см. [12]). В подавляющем большинстве известных автору рассматриваниях данного вопроса используются статические материальные уравнения  $\vec{D} = \epsilon_0 \epsilon \vec{E}$ ,  $\vec{B} = \mu_0 \mu \vec{H}$ . Основное возражение против ТЭИ Минковского состоит в том, что он несимметричен и для него не выполняется закон сохранения импульса [4]. В обеих формах ТЭИ компонента  $T_{44}$  определена как в статике, т.е. в виде  $T_{44} = u = (\epsilon_0 \epsilon \vec{E}^2 + \mu_0 \mu \vec{H}^2) / 2$ , тогда как в динамике имеется взаимодействие поля с веществом и наоборот, т.е. следует использовать выражения типа (56) [15, 16]. Не понятно, почему плотность энергии в среде должна зависеть от параметров среды, а

плотность импульса  $\vec{g}^A$  – нет, и быть такой же, как в вакууме. Ясно, что взаимодействие между полем и веществом может приводить к асимметрии ТЭИ поля и вещества [13]. Выше было показано, что плотность импульса зависит от временного процесса и определяется его предысторией. О конкретной форме ТЭИ поля в виде алгебраических соотношений, по-видимому, можно говорить только для гармонических процессов в средах с определенными законами дисперсии, когда эти гармонические процессы получены как предельный переход от соответствующих квазимонохроматических процессов создания поля, при котором усредненные за период компоненты  $\hat{T}$  выходят на стационарный уровень. Полный ТЭИ системы поле–вещество при произвольной дисперсии (временной и пространственной) пока не построен, и вопрос о требовании для него симметричности открыт. Если же тензоры поля и вещества симметричны по отдельности, это означает отсутствие взаимодействия между веществом и полем в смысле обмена энергией и импульсом.

В качестве примера покажем выполнение закона сохранения импульса для формы Минковского. Пусть плоский длинный с длиной  $l$  квазистационарный (квазимонохроматический) цуг или волновой пакет с прямоугольной огибающей и несущей частотой  $\omega$  (импульс в смысле нестационарной волны) в момент  $t_0 = 0$  подошел к границе слоя толщины  $d$  с постоянными проницаемостями  $\epsilon$ ,  $\mu$ . Квазистационарность означает, что  $l \gg \lambda_0 = 2\pi / k_0 = 2\pi c / \omega$ . Она нужна, чтобы можно было использовать квазимонохроматические значения величин. Длина волны в слое  $\lambda = \lambda_0 / n$ ,  $n = \sqrt{\epsilon\mu}$ . Пусть для простоты толщина  $d$  подобрана так, что на момент  $t_1$  весь цуг вошел в слой и полностью его заполнил. Поскольку скорость движения есть  $v_p = c/n$ , для этого времени получаем  $t_1 = d/v_p = nd/c$ . Для того чтобы это имело место, должно выполняться  $l = nd$ . Большую длину  $l$  для про-



стоты подберем так, чтобы на ней укладывалось целое  $M$  число длин волн:  $l = M\lambda_0$ ,  $M \gg 1$ , при этом  $d/\lambda = M$ . Имеется и отраженный импульс. На момент  $t_1$  он расположен в области  $-l \leq z \leq 0$  так же, как падающий импульс в момент  $t_0$ . Коэффициент отражения по электрическому полю нормально падающей плоской монохроматической волны от границы раздела есть  $R = (\rho_0 - 1)/(\rho_0 + 1)$ , где  $\rho_0 = \sqrt{\mu/\varepsilon}$  – нормированный к  $Z_0$  импеданс, а для коэффициента прохождения внутрь пластины имеем  $T = 2\rho_0/(\rho_0 + 1)$ . Падающая волна имеет вид (41) при  $\alpha = 0$ ,  $\varphi = 0$ . Далее будем рассматривать баланс для единичной площадки и средних значений плотностей импульсов. На момент  $t_0$  поле имело импульс  $G_0^M = g^M l = \langle S_0 \rangle l / c^2$  в направлении оси  $z$ , а импульс пластины был нулевым. Рассмотрим сначала идеальный случай согласованной пластины (stealth)  $\varepsilon = \mu$ . В этом случае отражения нет, и импульс пластине не передается. При интегрировании плотности импульса по  $z$  или после усреднения за период возникает множитель  $1/2$  у амплитудных значений. Соответственно  $g^M = Z_0 n^2 E_0^2 / (2c^2) = \langle u \rangle n / c$ , где  $\langle u \rangle = \varepsilon_0 \varepsilon E_0^2 / 2$ . В момент полного заполнения пластины  $t_1 = l/c$  импульс поля  $G_1^M = g^M d = dn^2 \langle S_1 \rangle / c^2 = dn \langle u_1 \rangle / c = nG_0^M$ . После того как цуг полностью вышел из пластины, т.е. в момент  $t_2 = 2t_1$  и после, вновь имеем  $G_1^M = G_0^M$ . Импульс Минковского как бы не сохраняется, поскольку пластина неподвижна. Такое «несохранение импульса в форме Минковского» выдвигается в [4] как основной аргумент против него и в пользу импульса Абрагама. Посмотрим, что будет с импульсом Абрагама. При  $t_0$  имеем:  $G_0^A = G_0^M = \langle S_0 \rangle l / c^2 = \langle u_0 \rangle l / c$ , а при  $t_1$  соответственно  $G_1^A = \langle S_1 \rangle d / c^2 = \langle u_0 \rangle d / c = G_1^M / n^2$ . Получили парадокс: ни импульс

Минковского, ни импульс Абрагама не сохраняются, при этом для первого сначала возникает, а затем исчезает дополнительный импульс, а для второго сначала часть импульса пропадает, а затем появляется вновь!!! В чем же дело? Для разрешения парадокса необходимо найти, где произошла подмена понятий или неправомерное их использование. В [4] аналогичные рассуждения приведены для цуга и пластины с  $\mu = 1$ ,  $n = \sqrt{\varepsilon}$ , причем предполагается, что пластина идеально согласованная с помощью просветляющего слоя. Идеально согласовать пластину нельзя, но для монохроматического процесса получить достаточно малый коэффициент отражения путем создания многослойного или неоднородного просветляющего слоя можно. Естественно, этот слой надо учитывать. Но у нас цуг с прямоугольными фронтами, имеющий все частоты, а процесс нестационарный!!! Конечно, основная спектральная интенсивность сосредоточена вблизи несущей частоты, но, тем не менее, пластина получит некоторый малый импульс. В случае  $\varepsilon = \mu$  отражения нет, но это условие может быть выполнено для дисперсивных искусственных сред в достаточно узкой полосе, при этом спектральные проницаемости всегда комплексные. Это означает, что реальная пластина из stealth материала также получит некий импульс. Далее мы от этого абстрагируемся. Заметим, что формула (1.13) из [4], якобы показывающая сохранение импульса Абрагама, на самом деле показывает его нарушение:  $G = e / nc$ . В ней  $e$  – полная энергия цуга. Она одина при любом его положении. В вакууме  $e = \langle u_0 \rangle l$ . В среде  $e = \langle u \rangle d$ , при этом, поскольку цуг сжимается в  $n$  раз,  $\langle u \rangle = n \langle u_0 \rangle$ , что непосредственно видно из (41). При этом полный импульс в обоих случаях есть  $G = e / c$ .

Так как же разрешить парадокс? Ясно, что для нашего случая в отсутствие силы Лоренца локальное уравнение баланса имеет вид

$$\partial u / \partial z = -\partial g^M / \partial t, \quad (57)$$

что следует из (41). Если выполняется локальный баланс, выполняется и глобальный (интегральный) баланс и его просто надо честно вычислить. Для этого заметим, что  $u(z, t)$  имеет скачки  $u_0(0, t)(n-1)$  при  $0 \leq t \leq t_1$  и  $u_0(d, t)(1-n)$  при  $t_1 \leq t \leq t_2$ . Соответственно в (57) появляются в указанные промежутки члены с дельта-функциями  $u_0(0, t)(n-1)\delta(z)$  и  $U_0(d, t)(1-n)\delta(z-d)$ .

Импульс в момент  $t_1$  следует вычислять так:

$$G_1^M = - \int_0^{t_1} \left[ \int_0^d u(z, t) dz + u_0(0, t)(n-1) \right] dt = \\ = nG_0^M - (n-1)G_0^M = G_0^M.$$

Указанный вывод о несохранении  $G^M$  в [4] основан на постоянстве скорости центра масс системы неподвижная пластина – движущийся электромагнитный цуг (фотон). Первая не обладает скоростью, а движущийся в одном направлении цуг (фотон) не обладает массой [40, 41]! Между тем вводится «масса фотона» из соотношения  $m = e/c^2$  и его импульс в виде  $G = mv_p = e/(cn)$ ! Говоря о цуге или фотоне в среде, следует иметь в виду, что плотность энергии возрастает в  $n$  раз, следовательно, возрастает и число фотонов, а скорость переноса энергии падает. Это значит, что в пластине имеются фотоны обоих направлений, движущиеся со скоростью  $c$  между актами элементарных взаимодействий с частицами вещества, а фотон с импульсом  $n\hbar\omega/c$  есть, по сути, квазичастица [3]. Полный интерференционный результат для макроскопической волны с большой энергией, в которой содержится много фотонов, есть следствие теоремы погашения, согласно которой волна в среде идет в прямом направлении с фазовой скоростью. В общем случае отражения от непоглощающей пластины при  $t > t_2$  имеем баланс энергии:  $1 = |R|^2 + |T|^2$ . Здесь введены полный коэффициент отражения  $R = (\rho_0^2 - 1) \tan(\theta) / ((\rho_0^2 + 1) \tan(\theta) - 2j\rho_0)$  и коэффициент прохождения  $T = (\cos(\theta) + j \sin(\theta)(Z^2 + 1) / (2Z))^{-1}$ ,

$\theta = \gamma l = 2\pi l / \lambda_0$ , поэтому отраженный импульс есть  $G_r^M = e|R|^2/c$ , а прошедший – соответственно  $G_t^M = e|T|^2/c$ , т.е. пластине передается импульс  $e(1 + |R|^2 - |T|^2)/c$ . Если  $R = 0$ , то  $T = 1$ , и пластина неподвижна. Если же  $\sigma \neq 0$ , то величина  $\gamma$  комплексная, баланс энергии такой:  $|R|^2 + |T|^2 < 1$ , и часть импульса внутри слоя передается ему. При очень малом  $|R|$  и большой толщине весь проходящий импульс передается пластине. При  $\sigma \rightarrow \infty$  имеем  $R \rightarrow -1$ ,  $T \rightarrow 0$ , и пластина получает удвоенный импульс. Здесь мы опять применили стационарные формулы к квазистационарным процессам вместо решения более сложной нестационарной задачи, что оправдано для длинного цуга. В стационарном же случае имеет смысл говорить лишь о давлении  $\langle u \rangle (1 + |R|^2 - |T|^2) / c$  или о переданном в секунду импульсе, при этом формулы уже точные. Что касается импульса  $G^A$ , то в рассмотренном случае он сохраняется, только если к нему прибавлен импульс, обусловленный силой Абрагама. Указанная сила для недиспергирующей среды возникает только при входе и выходе цуга из пластины [5]. Именно необходимость дополнения  $G^A$  неким импульсом для выполнения закона сохранения является основным аргументом против него. Другой аргумент против импульса Абрагама в среде состоит в том, что сила Абрагама точно не равна силам Лоренца, действующим на токи поляризации вещества.

Рассмотрим в качестве примера два атома с массами  $m_1$  и  $m_2$ , неподвижных в лабораторной системе при  $t < 0$ . Энергия системы имеет вид  $e = e_1 + e_2$ ,  $e_1 = m_1 c^2$ ,  $e_2 = m_2 c^2$ . Первый атом находится в возбужденном состоянии с энергией возбуждения  $\hbar\omega$  и расположен в начале координат, а второй – в точке  $(0, 0, z)$ ,  $z > 0$ . Возбужденный атом неподвижен (как квантовая частица) в



том смысле, что его волновая функция не зависит от времени (точнее, имеет множитель  $\exp(-ie_1 t / \hbar)$ ,  $i = -j$ ). При этом вероятность обнаружения атома в начале координат максимальна и выполняется соотношение неопределенности  $\Delta z \Delta p_z \geq \hbar/2$ . Это же относится и ко второму атому в точке  $z$ . Поскольку взаимодействия нет, полная волновая функция есть произведение волновых функций атомов и не зависит от времени до момента  $t_0 = 0$ . В момент  $t_0 = 0$  первый атом испускает фотон с энергией  $\hbar\omega$  и импульсом  $\vec{p} = \vec{z}_0 p = \vec{z}_0 \hbar\omega / c$ . Атом получает импульс  $-\vec{p}$  и движется влево, что означает появление зависимости его волновой функции от времени. Указанная волновая функция приобретает вид волнового пакета в виде собственной функции оператора импульса. Этот волновой пакет смещается влево со скоростью, удовлетворяющей соотношению  $(v_1/c) / \sqrt{1 - (v_1/c)^2} = p / (m'_1 c)$ , при этом  $m'_1 = \sqrt{(e_1 - \hbar\omega)^2 / c^4 - \vec{p}^2 / c^2} = m_1 \sqrt{1 - 2p / (m_1 c)}$  есть изменение массы в результате взаимодействия. Величина энергии возбуждения атома  $\hbar\omega$  играет роль внутренней неэлектромагнитной энергии, создающей поле. Пусть в момент  $t_1$  фотон в точке  $z = ct_1$  поглощается другим атомом. Указанный атом получает импульс  $\vec{p}$ , переходит в возбужденное состояние с энергией  $e_2 + \hbar\omega = m'_2 c^2 / \sqrt{1 - (v_2/c)^2}$ , при этом изменяется его масса:  $m'_2 = m_2 \sqrt{1 + 2p / (m_2 c)}$ . Скорость атома также определяется из соотношения  $(v_2/c) / \sqrt{1 - (v_2/c)^2} = p / (m'_2 c)$ . До момента  $t_0$  система двух невзаимодействующих атомов имела нулевой импульс, массу  $m_1 + m_2$  и энергию  $(m_1 + m_2)c^2$ . В промежутке  $t_0 < t < t_1$  масса атомов равна  $m'_1 + m_2$  и меньше их исходной массы, масса фотона равна нулю, но масса всей системы поле-вещество не

изменилась и равна  $m + m_0$ , поскольку её полный импульс есть нуль. При этом энергия и импульс сохраняются, а положение фотона не определено. Фотон локализуется (а поле исчезает) в момент поглощения (взаимодействия) в точке  $z$ . Масса всей системы опять равна  $m_1 + m_2$  и далее сохраняется, полный импульс также сохраняется и равен нулю. Сохраняется также и полная энергия  $e = m'_1 c^2 / \sqrt{1 - (v_1/c)^2} + m'_2 c^2 / \sqrt{1 - (v_2/c)^2}$ , при этом суммарная масса атомов равна  $m'_1 + m'_2 < m_1 + m_2$ . Образуется дефект массы, связанный с тем, что после взаимодействия атомы приобрели противоположные импульсы. Полные же величины энергии, массы и импульса всей замкнутой системы поле-вещество остались неизменными. Если  $2p / (m_i c) \ll 1$ ,  $i = 1, 2$ , то можно воспользоваться разложениями по малому параметру. Как видно, в первом порядке дефекта массы нет. Данный качественный пример, основанный на представлении нестационарного взаимодействия стационарными процессами с двумя точечными взаимодействиями, приведен для того, чтобы показать необходимость учета в балансе сторонних источников поля, в том числе и если оно имеет момент импульса. Он также полезен в связи с вышеприведенным примером из [4]. В сплошной среде распространение волны на микроскопическом уровне основано на подобных многочисленных актах взаимодействия, при этом скорость распространения зависит от времен жизни атомов в возбужденных состояниях, т.е. от характера рассеяния: упруго оно или нет, и насколько. В такой волне всегда есть преимущественный поток фотонов в направлении движения энергии, но возможны и обратные фотоны. На макроскопическом уровне волна создает токи поляризации, которые ее и поддерживают. Абсолютно упругое рассеяние фотонов на плазме свободных носителей происходит в пределе бесконечной проводимости, т.е. при  $\omega_p^2 / \omega_c \rightarrow \infty$ , или при  $\omega \rightarrow 0$ , что соответ-



ствуует отражению от идеальной электрической стенки. Из изложенного следует, что процессы обмена фотонами всегда, по крайней мере, квазистационарные. Поэтому монохроматическая волна есть в некотором роде удобная абстракция, которая принципиально в эксперименте получена быть не может.

В заключение заметим, что реальные законы дисперсии конкретных веществ весьма сложны. Необходимо рассматривать внутреннюю молекулярную или кристаллическую структуру, внутреннее поле, несколько (или несколько десятков) собственных резонансных частот, имеющуюся пространственную дисперсию, а также применять нестационарный подход. Это означает, что в ТЭИ системе поля и вещества не определяется только значениями полей в данный момент, т.е. должна быть учтена предыстория создания поля источниками и соответственно предыстория воздействия на вещество сил источников и поля. Если источников в рассматриваемом объеме нет, такая предыстория должна быть учтена с момента вхождения созданного внешними источниками поля в указанный объем. Поэтому следует брать плотность импульса в виде  $\vec{g}^M = \vec{D} \times \vec{B}$  с учетом вычислений (1), а для плотности энергии вычислять соотношения типа (56) [15, 16]. При этом полный ТЭИ системы поле–вещество несимметричен. В случае отсутствия дисперсии он переходит в ТЭИ Минковского. Для плоской волны величина  $\hat{\partial}_z u$ , определяемая соотношением (56), не дает плотность потока импульса. В общем случае использование только материальных уравнений позволяет разделить импульс поля и вещества и соответствующие им потоки. Для плотности потока энергии такое разделение без решения уравнений движения вещества, по видимому, не имеет места. Это связано с тем, что плотность мощности токов поляризации в правой части (9), содержащая производные по времени, скорее, относится к полной плотности мощности, нежели является дивергенцией некоего вектора. К тому же указанный вектор определяется неоднозначно. Кроме этого энергию взаимодействующей системы поле–вещество в общем случае не-

возможно разделить на энергию поля и вещества без учета энергии взаимодействия. Основное возражение против ТЭИ Минковского состоит в том, что он несимметричен. Требование симметричности возникает из требования однозначности определения ТЭИ и условия, что тензор момента следует определять через ТЭИ по стандартным формулам, при этом используется связь компонент импульса с компонентами ТЭИ такая, как для поля в вакууме ([52], с.107). Очевидно, такой ТЭИ есть тензор поля в вакууме, совпадающий с тензором Абрагама. В нашем случае нестационарного возбуждения в среде ТЭИ определяется однозначно, и дополнительного условия не требуется. Тензор же момента импульса системы поле–вещество следует определять отдельно. При этом необходимо учитывать, обладало или нет вещество моментом импульса до создания поля, а также момент, передаваемый сторонним источникам при генерации поля.

Приведенные выражения для ТЭИ зависят от формы материальных уравнений. Если материальные соотношения взять в другой форме, получим другие соотношения. В последнее время встает вопрос о ТЭИ и о соответствующих плотностях (и скоростях) в искусственных средах с пространственной дисперсией, включая и бианизотропные левые среды (см., например, подобную неудачную попытку в [53]), модели которых весьма сложны. Однако в ряде случаев законы дисперсии в определенных частотных областях можно описать более простыми соотношениями (моделями), что и использовано в данной работе.

#### Список литературы

1. *Minkowski H.* Die Grungleichungen für die electromagnetischen Vorgänge in bewegten Körpern // Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, 1908. P.111.
2. *Abraham M.* Zur Electrodynamik bewegter Körper // Rendiconti Circolo Matematico di Palermo. 1909. Vol.28. P.1–28.
3. *Гинзбург В.Л.* О законах сохранения энергии и импульса при излучении электромагнитных волн (фотонов) в среде и о тензоре энергии-импульса в макроскопической электродинамике // УФН. 1973. Т.110, вып.2. С.309–319.
4. *Скобельцын Д.В.* О тензоре импульс-энергии электромагнитного поля // УФН. 1973. Т.110, вып.2. С.253–292.
5. *Гинзбург В.Л., Угаров В.А.* Несколько замечаний о силах и тензоре энергии-импульса в макроскопической электродинамике // УФН. 1976. Т.118, вып.1. С.175–188.



6. *Brevik I.* Electromagnetic energy-momentum tensor within material media. II. Discussion on various tensor forms // *Mat. Phys. Med. Dan. Vid. Selsc.* 1970. Vol.37, №13. P.1–79.
7. *Groot S. de, Suttorp L.* The relativistic energy-momentum tensor in polarized media: IV. The macroscopic material energy-momentum tensor // *Physica.* 1968. Vol.38. P.84.
8. *Groot C.P. de, Samton L.G.* Электродинамика. М.: Наука, 1982. 530 с.
9. *Скобельцын Д.В.* Парадоксы квантовой теории эффектов Вавилова–Черенкова и Доплера // *УФН.* 1977. Т.122. С.295.
10. *Гинзбург В.Л.* Замечания к статье Д.В. Скобельцына «Парадоксы квантовой теории эффектов Вавилова–Черенкова и Доплера» // *УФН.* 1977. Т.122. С.325.
11. *Leonhardt U.* Momentum in an uncertain light // *Nature.* 2006. Vol.444. P.823–824.
12. *Pfeifer R.N., Nieminen T.A., Heckenberg N.R., Rubinsztein-Dunlop H.* // *Rev. Mod. Phys.* 2007. Vol.79. P.197.
13. *Obukhov Yu.N.* Electromagnetic energy and momentum in moving media // *Annalen der Physik.* 2008. №8. P.1–22.
14. *Давидович М.В.* О парадоксе Хартмана, туннелировании электромагнитных волн и сверхсветовых скоростях // *УФН.* 2009. Т.179, №4. С.443–446.
15. *Ахиезер А.И., Ахиезер И.А.* Электромагнетизм и электромагнитные волны. М.: Высш. шк., 1985. 504 с.
16. *Давидович М.В.* О плотности электромагнитной энергии и ее скорости в среде с аномальной положительной дисперсией // *Письма в ЖТФ.* 2006. Т.32, вып.22. С.53–63.
17. *Виноградов А.П.* К вопросу о материальных уравнений в электродинамике // *УФН.* 2002. Т.172, №3. С.363–370.
18. *Ландау Л.Д., Лившиц Е.М.* Электродинамика сплошных сред. М.: Наука, 1982. 624.
19. *Фейнман Р., Лейтон Р., Сэндс М.* Фейнмановские лекции по физике. Электродинамика. М.: Мир, 1966. Т.6. 260 с.
20. *Гайтлер В.* Квантовая теория излучения. М.: Иностран. лит., 1956. 492 с.
21. *Умов Н.А.* Теория простых сред и её приложение к вопросу основных законов электростатических и электродинамических взаимодействий. Одесса, 1873.
22. *Окунь Л.Б.* Понятие массы (Масса, энергия, относительность) // *УФН.* 1989. Т.158, вып.3. С.511–530.
23. *Окунь Л.Б.* О письме Р.И. Храпко «Что есть масса?» // *УФН.* 2000. Т.170, №12. С.1366–1371.
24. *Котельников И.А.* О затухании в волноводе // *ЖТФ.* 2004. Т.74, вып.9. С.91–96.
25. *Ordal M.A., Long L.L., Bell R.J. et al.* Optical properties of the metals Al, Co, Cu, Au, Fe, Pb, Ni, Pd, Pt, Ag, Ti, and W in the infrared and far infrared // *Appl. Opt.* 1983. Vol.22, №7. P.1099–1120.
26. *Марков Г.Т., Чаплин А.Ф.* Возбуждение электромагнитных волн. М.: Радио и связь, 1983. 296 с.
27. *Зильбергейт А.С., Копилевич Ю.И.* О скорости переноса энергии волнами // *ЖТФ.* 1980. Т.50, №2. С.241–251.
28. *Зильбергейт А.С., Копилевич Ю.И.* О скорости переноса энергии волнами // *ЖТФ.* 1980. Т.50, №3. С.449–460.
29. *Гуреев А.В.* О скорости переноса энергии электромагнитными волнами в регулярном экранированном волноводе // *ЖТФ.* 1990. Т.61, №1. С.23–28.
30. *Силин Р.А.* Периодические волноводы. М.: Фазис, 2002. 438 с.
31. *Мандельштам Л.И.* Лекции по оптике, теории относительности и квантовой механике. М.: Наука, 1972. 438 с.
32. *Бхатнагар П.* Нелинейные волны в одномерных дисперсных средах. М.: Мир, 1983. 136 с.
33. *Гинзбург В.Л.* Распространение электромагнитных волн в плазме. М.: Физматлит, 1960. 550 с.
34. *Вайнштейн Л.А.* Электромагнитные волны. М.: Радио и связь, 1988. 440 с.
35. *Гольдштейн Л.Д., Зернов Н.В.* Электромагнитные поля и волны. М.: Сов. радио, 1971. 654 с.
36. *Виноградова М.Б., Руденко О.В., Сухоруков А.П.* Теория волн. М.: Наука, 1979. 384 с.
37. *Вайнштейн Л.А.* Распространение импульсов // *УФН.* 1976. Т.118, вып.2. С.339–367.
38. *Вайнштейн Л.А., Вацман Д.Е.* Разделение частот в теории колебаний и волн. М.: Наука, 1983. 288 с.
39. *Stratton J.A.* Electromagnetic theory. N.Y.; L.: McGraw-Hill Book Company, Inc., 1941.
40. *Schulz-DuBois E.O.* Energy Transport Velocity of Electromagnetic Propagation in Dispersive Media // *Proc. of the IEEE.* 1969. Vol.57, №10. P.1748–1757.
41. *Островский Л.А., Потанов А.И.* Введение в теорию модулированных волн. М.: Физматлит, 2003. 400 с.
42. *Полевой В.Г., Рытов С.М.* О четырехмерной групповой скорости // *УФН.* 1978. Т.125, вып.3. С.540–565.
43. *Рытов С.М.* Некоторые теоремы о групповой скорости электромагнитных волн // *ЖЭТФ.* 1947. Т.17(10). С.930.
44. *Lighthill M.J.* Group Velocity // *J. Inst. Math. and its Appl.* 1965. Vol.1. P.1–28.
45. *Biot V.A.* General theorems on the equivalence of group velocity and energy transport // *Phys. Rev.* 1957. Vol.105. P.1129–1137.
46. *Ольвер Ф.* Введение в асимптотические методы и специальные функции. М.: Наука, 1973. 376 с.
47. *Mayergoyz I.D.* Some remarks concerning electromagnetic potentials // *IEEE Trans.* 1993. Vol.Magn.-29, №2. P.1301–1305.
48. *Hinds E.A., Barnett S.M.* Momentum exchange between light and single atom: Abraham or Minkowski? // *Phys. Rev. Lett.* 2009. Vol.102(5). P.050403.
49. *Marklund M.* Radiation transport in diffractive media // *J. Phys. A, Math. Gen.* 2005. Vol.38, №19. P.4265–4273.
50. *Garrison J.C., Chiao R.Y.* Canonical and kinetic forms of the electromagnetic momentum in ad hoc quantization scheme for a dispersive dielectric // *Phys. Rev. A.* 2004. Vol.70, №5. P.053826.
51. *Leonhardt U.* Energy-momentum balance in quantum dielectrics // *Phys. Rev. A.* 2006. Vol.73. P.032108.
52. *Brevik I.* Electromagnetic energy-momentum tensor within material media. I. Minkowski tensor // *Mat. Phys. Medd. Dan. Vid. Selsc.* 1970. Vol.37, №11. P.1–52.
53. *Brevik I.* Electromagnetic energy-momentum tensor within material media // *Mat. Phys. Medd. Dan. Vid. Selsc.* 1970. Vol.37, №13. P.1–52.
54. *Ландау Л.Д., Лившиц Е.М.* Теория поля. М.: Наука, 1973. 504 с.
55. *Веселаго В.Г.* Перенос энергии, импульса и массы при распространении электромагнитной волны в среде с отрицательным преломлением // *УФН.* 2009. Т.179, №6. С.689–694.



## СВЕДЕНИЯ ОБ АВТОРАХ

**Бабков Лев Михайлович** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра теоретической и математической физики, профессор. E-mail: babkov@sgu.ru

**Вакас-Жакс Паулино** – Университет Леона, центр оптических исследований, Гуанахиато, Мексика, аспирант. E-mail: pvj@cio.mx

**Давидович Михаил Владимирович** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра радиотехники и электродинамики, профессор. E-mail: DavidovichMV@info.sgu.ru

**Давыдова Надежда Александровна** – Институт физики НАН Украины, Киев, старший научный сотрудник. E-mail: davydova@iop.kiev.ua

**Джалмухамбетова Елена Азатуллаевна** – Астраханский государственный университет, кафедра общей физики, доцент. E-mail: alenna@list.ru

**Зайко Юрий Николаевич** – Поволжская академия государственной службы им. П.А. Столыпина, Саратов, кафедра прикладной информатики и информационных технологий, доцент. E-mail: zytnick@rambler.ru

**Королевич Майя Васильевна** – Институт физики им. Б.И. Степанова НАН Беларуси, Минск, ведущий научный сотрудник. E-mail: korolevi@dragon.bas-net.by

**Моисейкина Елена Александровна** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра оптики и биомедицинской физики, студентка. E-mail: moisejkinaelena2007@rambler.ru

**Назимов Алексей Игоревич** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра радиофизики и нелинейной динамики, студент. E-mail: pavlov.lesha@gmail.com

**Павлов Алексей Николаевич** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра радиофизики и нелинейной динамики, доцент. E-mail: pavlov.lesha@gmail.com

**Павлова Ольга Николаевна** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра радиофизики и нелинейной динамики, ассистент. E-mail: pavlova\_olga@yahoo.com

**Паез Гонзало** – Университет Леона, центр оптических исследований, Гуанахиато, Мексика, профессор. E-mail: gpaez@cio.mx

**Постнов Дмитрий Энгелевич** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра радиофизики и нелинейной динамики, профессор. E-mail: postnov@info.sgu.ru

**Рябухо Владимир Петрович** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра оптики и биомедицинской физики, профессор. E-mail: rvr@sgu.ru

**Стройник Мария** – Университет Леона, центр оптических исследований, Гуанахиато, Мексика, профессор. E-mail: mstrojnik@aol.com

**Тупицын Анатолий Николаевич** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра радиофизики и нелинейной динамики, аспирант. E-mail: tupitsin-anatoliy@yandex.ru

**Тучин Валерий Викторович** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра оптики и биомедицинской физики, заведующий кафедрой, профессор. E-mail: tuchin@sgu.ru

**Успенский Кирилл Евгеньевич** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра теоретической и математической физики, аспирант. E-mail: babkov@sgu.ru

**Элькин Михаил Давыдович** – Саратовский государственный университет им. Н.Г. Чернышевского, кафедра прикладной оптики и спектроскопии, профессор. E-mail: elkinmd@mail.ru





## INFORMATION ABOUT THE AUTHORS

**Babkov Lev Mikhailovich** – Saratov State University, chair of theoretical and mathematical physics, professor. E-mail: babkov@sgu.ru

**Vacas-Jacque Paulino** – Centro de Investigaciones en Optica, 37000, Leon, Guanajuato, Mexico, post-graduate. E-mail: pvj@cio.mx

**Davidovich Mikhail Vladimirovich** – Saratov State University, chair of radio engineering and electrodynamics, professor. E-mail: DavidovichMV@info.sgu.ru

**Davydova Nadezda Alexandrovna** – Institute of physic NAS Ukraine, Kiev, senior research. E-mail: davydova@iop.kiev.ua

**Dzalmuhambetova Elena Azatullaevna** – Astrakhan State University, chair of general physics, reader. E-mail: alenna@list.ru

**Zayko Yuri Nikolaevich** – Stolypin's Volga Regional Academy of State Service, Saratov, chair of applied informatics and information technologies in control, reader. E-mail: zymick@rambler.ru

**Korolevich Maiya Vasil'evna** – B.I. Stepanov Institute of physic NAS Belarus, Minsk, leader research. E-mail: korolevi@dragon.bas-net.by

**Moiseykina Elena Alexandrovna** – Saratov State University, chair of optics and biomedical physics, student. E-mail: moisejkinaelena2007@rambler.ru

**Nazimov Alexyi Igorevich** – Saratov State University, chair of radiophysics and nonlinear dynamics, student. E-mail: pavlov.lesha@gmail.com

**Pavlov Alexyi Nikolaevich** Saratov State University, chair of radiophysics and nonlinear dynamics, reader. E-mail: pavlov.lesha@gmail.com

**Pavlova Olga Nikolaevna** – Saratov State University, chair of radiophysics and nonlinear dynamics, assistant. E-mail: pavlova\_olga@yahoo.com

**Paez Gonzalo** – Centro de Investigaciones en Optica, 37000, Leon, Guanajuato, Mexico, professor. E-mail: gpaez@cio.mx

**Postnov Dmitriy Engelevich** – Saratov State University, chair of radiophysics and nonlinear dynamics, professor. E-mail: postnov@info.sgu.ru

**Ryabukho Vladimir Petrovich** – Saratov State University, chair of optics and biomedical physics, professor. E-mail: rvp@sgu.ru

**Strojnuk Marija** – Centro de Investigaciones en Optica, 37000, Leon, Guanajuato, Mexico, professor. E-mail: mstrojnuk@aol.com

**Tupitsyn Anatoly Nikolaevich** – Saratov State University, chair of radiophysics and nonlinear dynamics, post-graduate. E-mail: tupitsin-anatoliy@yandex.ru

**Tuchin Valery Victorovich** – Saratov State University, head of chair of optics and biomedical physics, professor. E-mail: tuchin@sgu.ru

**Uspenskiy Kirill Evgen'evich** – Saratov State University, chair of theoretical and mathematical physics, post-graduate. E-mail: babkov@sgu.ru

**El'kin Mikhail Davydovich** – Saratov State University, chair of applied optics and spectroscopy, professor. E-mail: elkinmd@mail.ru