E10-2015-67

N. B. Bogdanova [1,2,*], S. T. Todorov [1,**], G. A. Ososkov [2,***]

USE OF ORTHONORMAL POLYNOMIAL EXPANSION METHOD TO THE DESCRIPTION OF THE ENERGY SPECTRA OF BIOLOGICAL LIQUIDS

[1] Institute for Nuclear Research and Nuclear Energy of the Bulgarian Academy of Sciences, Sofia

[2] Joint Institute for Nuclear Research, Dubna

[*] E-mail: nibogd@inrne.bas.bg

[**] E-mail: todorov_st@yahoo.com

[***] E-mail: ososkov@jinr.ru

Богданова Н. Б., Тодоров С. Т., Ососков Г. А. E10-2015-67
Использование метода разложения по ортонормированным полиномам
для описания энергетических спектров биологических жидкостей

Метод разложения по ортонормированным полиномам применяется для описания данных, полученных методом энергетических спектров для жидких биомасс, выращенных с использованием гербицидов. Так как биомасса биологических объектов содержит жидкости, состоящие в основном из воды, метод водных спектров может быть применен и в этом случае. Рассматриваются также похожие данные, полученные на контрольном образце, состоящем из жидкости пшеницы, выращенной без применения гербицидов. Для описания спектров был использован метод ортонормированных полиномов, основанный на полной (эффективной) дисперсии, учитывающей ошибки по обеим (зависимой и независимой) переменным. Были применены специальные критерии оценки оптимальной степени полинома и числа итераций. Результаты расчета хорошо описывают экспериментальные данные в форме, удобной для дальнейшего анализа в области теоретической экологии.

Работа выполнена в Лаборатории информационных технологий ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна, 2015

Bogdanova N. B., Todorov S. T., Ososkov G. A. E10-2015-67
Use of Orthonormal Polynomial Expansion Method
to the Description of the Energy Spectra of Biological Liquids

Orthonormal polynomial expansion method (OPEM) is applied to the data obtained by the method of energy spectra to the liquid of the biomass of wheat in the case when herbicides are used. Since the biomass of a biological object contains liquid composed mainly of water, the method of water spectra is applicable to this case as well. For comparison, the similar data obtained from control sample consisting of wheat liquid without the application of herbicides are shown. The total variance OPEM is involved including errors in both dependent and independent variables. Special criteria are used for evaluating the optimal polynomial degree and the number of iterations. The presented numerical results show good agreement with the experimental data. The developed analysis frame is of interest for future analysis in theoretical ecology.

The investigation has been performed at the Laboratory of Information Technologies, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna, 2015

1. EXPERIMENTAL DATA

The method of liquid spectra uses a liquid drop taken from a particular liquid probe (or liquid on which some treatment was applied) to measure the probe's state spectrum. For this purpose the drop is placed on a substrate (hostaphan). During the process of evaporation of the drop, one measures at equal time intervals the drop contact angle with the substrate. In this way one gets the following graph. On X-axis one has the contact angle values and on Y-axis — the numbers proportional to the frequency of measurements of the given contact angle lying within a fixed angular interval. Simultaneously with the probe measurement F, one measures a probe of similar but untreated sample (control sample) $F0$, see Fig. 1. The arithmetic difference between the two spectra is called "differential spectrum", which is independent of external influences on the spectrum of the probe. Various applications of this approach are described in [1–3].

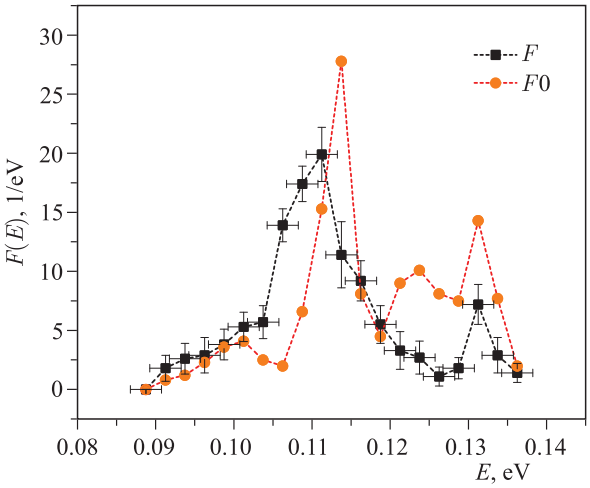

Fig. 1. Experimental data F with their errors in both variables and control curve $F0$

One can change the variable on the X-axis from angle variable to energy variable, according to Antonov's transformation given by the formula:

$$F(E) = F(\theta)/\sqrt{1-(1+bE)^2}, \tag{1}$$

where

$$b = I(1+\cos\theta_0)/\gamma. \tag{2}$$

Here $I = 5.0310^{18}m^{-2}$ is the density of the liquid molecules, γ is the surface tension, θ_0 is the initial contact angle. The graph obtained this way is referred to as energy spectrum of the sample. In these variables the X range of the obtained spectrum contains the value of the liquid hydrogen bond energy. The method of energy spectra is described in some details in [1]. The method is sensitive to discovering external influences on the sample [2].

The method of water spectra is applicable not only to water samples. Since the biomass of a biological object contains biological liquid composed mainly of water, the method of water spectra is applicable to this case as well. In our case we apply the OPEM to energy spectrum of wheat liquid treated with herbicides [3]. The wheat liquid is taken from the plant in its initial period of development.

2. MATHEMATICAL APPROACH

Here one defines new variance at i^th given point (E_i, F_i) using the expression [4,5]

$$S_i^2 = {\sigma_{F_i}}^2 + \left(\frac{\partial F_i}{\partial E_i}\right)^2 {\sigma_{E_i}}^2, \tag{3}$$

where σ_{F_i} and σ_{E_i} are the given standard deviations in both variables, (E_i, F_i). In the formula (3) Bevington's [4] proposal to combine both variable uncertainties and assign them to the dependent variable is used.

2.1. The Generalized OPEM. The basic recurrence for one-dimensional generation of orthonormal polynomials by Forsythe [6] $\{P_i^{(0)},\ i = 1, 2, \ldots\}$ and their derivatives $\{P_i^{(m)},\ m = 1, 2, \ldots\}$ in the OPEM is

$$P_{i+1}^{(m)}(E) = 1/\nu_{i+1}[(E-\mu_{i+1})P_i^{(m)}(E) - \\ - (1-\delta_{i0})\nu_i P_{i-1}^{(m)}(E) + mP_i^{(m-1)}(E)]. \tag{4}$$

The generalization of the Forsythe procedure in the one-dimensional case assumes arbitrary weights at every point, evaluating derivatives ($m > 0$) or integrals ($m < 0$) and normalizing polynomials. Here the normalization coefficient $1/\nu_i$ and the recurrence coefficients μ_i, ν_i are given as scalar products of the polynomials in the given data in M points [7]. We developed some features of

our algorithm [8–13]. One can generate $P_i^{(m)}(E)$ recursively. The polynomials satisfy the following orthogonality relations:

$$\sum_{i=1}^{M} w_i P_k^{(0)}(E_i) P_l^{(0)}(E_i) = \delta_{kl}$$

over the discrete point set $\{E_i,\ i = 1, 2, \dots M\}$, where $w_i = 1/(\sigma_{F_i}^2)$ are the corresponding weights. The approximation function $F^{(0)\mathrm{ap}}$ and its derivatives $F^{(m)ap}$ are constructed as follows with orthonormal coefficients a_k:

$$F^{(m)\mathrm{ap}}(E) = \sum_{k=0}^{L} a_k P_k^{(m)}(E). \tag{5}$$

The coefficient matrix in the least square method becomes an identity matrix, and due to orthogonality conditions the coefficients a_k are easily computed by

$$a_k = \sum_{i=1}^{M} F_i w_i P_k^{(m)}(E_i). \tag{6}$$

The approximation function F^{ap} is presented in orthonormal and usual (ordinary) expansion:

$$F^{\mathrm{ap}}(E) = \sum_{k=0}^{L} a_k P_k(E) = \sum_{k=0}^{L} c_k E^k. \tag{7}$$

Let us write polynomials in the ordinary basis:

$$P_k = \sum_{j=0}^{k} c_j^{(k)} E^j, \quad k = 0, \dots, L. \tag{8}$$

Here the superscript in $c_j^{(k)}$ indicates the order of polynomials. Then we have the relation:

$$c_j = \sum_{i=j}^{L} a_i c_j^{(i)}, \quad j = 0, 1, 2, \dots, L. \tag{9}$$

The knowledge of a_k enables one to calculate c_k with the help of the coefficients μ_i, ν_i from formula (5). They are explicitly constructed by recurrence [9, 11–13]. We remark that, for the sake of uniformity, all the calculations are carried out for E in $[-1, 1]$, i.e., after the input data E_i in the interval $[E_1, E_M]$, $E_1 \leqslant E_i \leqslant E_M$ are mapped onto the unit interval $[-1, 1]$. The inherited errors in usual coefficients are given by the inherited errors in orthogonal coefficients:

$$\Delta c_j = \left(\sum_{i=j}^{L} (c_j^{(i)})^2 \right)^{1/2} \Delta a_i, \quad j = 0, 1, 2, \dots, L. \tag{10}$$

And the inherited errors in orthonormal coefficients are expressed by

$$\Delta a_i = \left[\sum_{k=1}^{M} P_i^2(E_k) w_k (F_k - F_k^{\rm ap})^2 \right]^{1/2}, \ i = 0, \ldots, L. \tag{11}$$

It is worth noting the following advantages of the OPEM:
a) It avoids recomputing the coefficients in Eq. (6) for evaluating approximation with higher degree polynomials — the coefficients got at the lower-order polynomials remain unchanged.
b) It avoids the usual matrix inversion to obtain the solution. This shortens the computing time and the memory storage. Case study solutions point to a significant decrease in the number of iterations required to reach a prescribed numerical precision.

The development now is carried out to solve the given approximation task with errors in both variables. Two criteria are used here to select the optimum series length in Eq. (5).

2.1.1. First Criterion. (i) Here one neglects the errors in E variable, the graph of the fitting curve lies inside the "old" error corridor $[F - \sigma, F + \sigma]$.
(ii) After calculating the derivatives at any point E_i using Eqs. (4) and (5), the fitting curve has to lie inside the total error corridor $[F - S, F + S]$.

2.1.2. Second Criterion. We extend the above algorithm to include S_i^2 in the OPEM in two stages:

(i) The following χ^2 is minimized:

$$\chi^2 = \sum_{i=1}^{M} w_i [F^{\rm ap}(E_i) - F(E_i)]^2 / (M - L - 1),$$

where the weights are $w_i = 1/\sigma_{F_i}^2$.
(ii) The next approximation is calculated with the weight function $w_i = 1/S_i^2$.

The results of calculations in (i) give the first approximation. The procedure is iterative and the result of the $k^{\rm it}$-th iteration, $k^{\rm it} > 1$, is called below the $k^{\rm it}$-th approximation.

We have to note the third advantage:
c) At every iteration step the algorithm chooses automatically the optimal degree using the two criteria of above simultaneously. The preference is given to the first criterion, and when it is satisfied, the search for the minimal chi-squared stops. Based on the above features, the algorithm selects the optimal solution for a given set $\{E, F\}$.

2.1.3. Criterion for Ordinary Expansion. After evaluating automatically an optimal number of polynomials in orthonormal expansion, we find the best result

of the usual expansion at every step of iteration by minimization of

$$\max|E_a^{\rm ap} - E_c^{\rm ap}| = \max{}_{i=1}^{M}|F_a^{\rm ap}(E_i) - F_c^{\rm ap}(E_i)|. \qquad (12)$$

Now the algorithm is called the total (effective) OPEM.

3. APPROXIMATION RESULTS

The iterations are carried out till the tenth one. The optimal number of polynomials is chosen between 2 and 7. Table 1 shows in every step iteration the optimal degree L with corresponding χ^2 and maximal absolute difference between orthonormal and usual approximating values $\max|E_a^{\rm ap} - E_c^{\rm ap}|$.

In Table 2 and Figs. 2, 3, the best results with the $L = 6$ and $k^{\rm it} = 6$ are given. The approximating values in two expansions for $k^{\rm it} = 6$ and $L = 6$ are shown explicitly in Table 2.

Table 1. **OPEM approximation results for energy spectra with every step approximation**

$k^{\rm it}$	2	3	4	5	6	7	8	9	10
$L(2 \div 7)$	7	6	6	6	6	6	6	6	6
$\sqrt{\chi^2} \cdot 10^{-1}$	8.68	8.19	7.74	7.59	7.56	7.56	7.56	7.56	7.56
$\max\|(F_a^{\rm ap} - F_c^{\rm ap})\|$	24.12	4.4	7.69	6.27	3.58	7.49	9.95	6.60	4.75

Table 2. **Experimental and OPEM approximating values of contact energy spectra**

No.	E	F	σ_E	σ_F	$F_a^{\rm ap,6,6}$	$F_c^{\rm ap,6,6}$	$\lvert F_a^{\rm ap,6,6} - F_c^{ap,6,6}\rvert$
1	0.091	1.8	0.002	1.1	2.2681	1.5791	0.6892
2	0.093	2.6	0.002	1.3	2.6079	1.7609	0.8478
3	0.096	2.9	0.002	1.5	2.3735	1.3030	1.0709
4	0.098	3.8	0.002	1.3	3.5273	2.6941	0.8329
5	0.101	5.3	0.002	1.25	6.3619	5.6700	0.6913
6	0.103	5.7	0.002	1.4	10.1895	8.8068	1.3820
7	0.106	13.9	0.002	1.4	13.9053	12.3100	1.5943
8	0.108	17.4	0.002	1.5	16.4263	14.4264	1.9999
9	0.111	19.9	0.002	2.3	17.0048	15.6354	1.3696
10	0.113	11.4	0.002	2.8	15.4168	14.4860	0.9312
11	0.116	9.2	0.002	1.7	12.1598	10.7609	1.2646
12	0.118	5.5	0.002	1.6	7.7174	6.7466	0.9710
13	0.121	3.3	0.002	1.6	3.7186	1.4054	2.3133
14	0.123	2.7	0.002	1.4	1.2793	1.0626	0.2168
15	0.126	1.1	0.002	0.8	1.2377	–0.4110	1.6493
16	0.128	1.8	0.002	0.9	3.4573	–0.1287	3.5869
17	0.131	7.2	0.002	1.7	6.1386	7.8103	1.6706
18	0.133	2.9	0.002	1.5	5.0068	2.0849	2.9220

In Fig. 2 we present the given data with errors in both variables and the obtained approximated values for $L = 6$ and $k^{\rm it} = 6$ with orthonormal polynomials. One can see good agreement for curves in more points, following a new corridor $[F - S, F + S]$. And in Fig. 3 we have compared the given curve and two approximated curves by two expansions. The good agreement between the three curves is obvious.

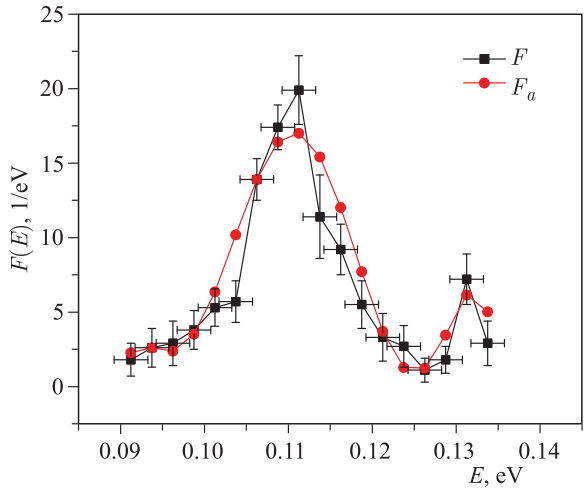

Fig. 2. Experimental data F with their errors and approximated by sixth degree's orthonormal coefficients F_a values

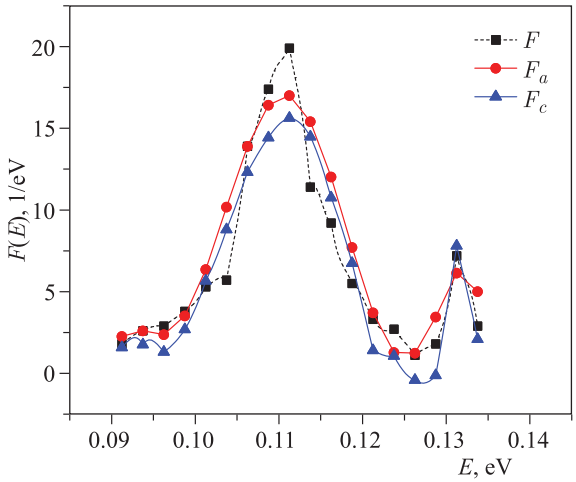

Fig. 3. Experimental data F with their orthonormal $F_a^{\rm ap,6,6}$ and usual $F_c^{\rm ap,6,6}$ expansions

CONCLUSIONS

• The results given in tables and figures demonstrate the appropriateness of the selection criteria following from orthonormal and ordinary expansion approximation.

• The best accuracy of fitting results from Tables 1 and 2 is got at $L = 6$ and $k^{\mathrm{it}} = 6$.

• The new fitting curves in both expansions are close enough to each other (see the last column in Table 2).

• The present version of the total OPEM approximation gives good results for further interpretations and comparisons in theoretical ecology.

Acknowledgements. The authors are grateful to Prof. Gh. Adam, Prof. I. V. Puzynin, and Dr. B. F. Kostenko for useful discussions.

REFERENCES

1. *L. Todorova, A. Antonov*, Compt. Rend. Acad. Bulg. Sci., 2000. V. 53, No. 7. P. 43–45.
2. *A. Antonov, L. Todorova*, Compt. Rend. Acad. Bulg. Sci., 1995. V. 48, No. 1. P. 21–24.
3. *St. T. Todorov*, Ecology and Health 2008, Proc. of VII Scientific-Technical Conference, with International Participation, House of Science and Technology. Plovdiv, P. 363–366.
4. *P. R. Bevington*, Data Reduction and Error Analysis for the Physical Sciences. McGraw-Hill, New York, 1969.
5. *G. Jones*, Preprint TRI-PP-92-31, A. 1992.
6. *G. Forsythe*, J. Soc. Ind. Appl. Math. 1957. V. 5, No. 2. P. 74–88.
7. *N. Bogdanova*, J. Comp. Appl. Math. 1986. V. 14, No. 3. P. 345–351.
8. *V. Gadjokov, N. Bogdanova*, JINR Commun. P11-12860. Dubna, 1979 (in Russian).
9. *N. Bogdanova*, JINR Commun. E11-98-3. Dubna, 1998.
10. *N. Bogdanova, St. Todorov*, IJMPC. 2001. V. 12, No. 1. P. 119–126.
11. *N. Bogdanova*, reported at BPU6 Conference, Istanbul, August 2006, in 2007 AIP proceedings, Eds. S. A. Cetin, I. Hikmet, P. 377.
12. *N. Bogdanova, S. Todorov*, Bull. PFUR, ser. Mathematics, Information, Physics, 2010. No. 3(2). P. 63–67.
13. *N. Bogdanova, S. Todorov*, Phys. Part. Nucl. Lett. 2015. V. 12, No. 3. P. 385–388.

Received on 29 July, 2015.

Редактор *Е. И. Крупко*

Подписано в печать 08.09.2015.
Формат 60 × 90/16. Бумага офсетная. Печать офсетная.
Усл. печ. л. 0,62. Уч.-изд. л. 0,88. Тираж 225 экз. Заказ № 58628.

Издательский отдел Объединенного института ядерных исследований
141980, г. Дубна, Московская обл., ул. Жолио-Кюри, 6.
E-mail: publish@jinr.ru
www.jinr.ru/publish/