Письма в ЭЧАЯ. 2015. Т. 12, № 3(194). С. 608–613

NEW APPROXIMATING RESULTS FOR DATA WITH ERRORS IN BOTH VARIABLES

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

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

We introduce new data from mineral water probe in Lenovo, Bulgaria, measured with errors in both variables. We apply our Orthonormal Polynomial Expansion Method (OPEM), based on the Forsythe recurrence formula, to describe the data in the new errors corridor. The development of OPEM gives the approximating curves and their derivatives in optimal orthonormal and usual expansions, including the errors in both variables, with special criteria.

Мы рассматриваем новые данные для пробы минеральной воды из источника в Леново (Болгария), измеренные с ошибками по обеим переменным. Для описания данных в новом коридоре ошибок применяется наш метод разложения по ортогональным полиномам, основанный на рекуррентной формуле Форсайта. Представлено дальнейшее развитие нашего численного метода, связанное с использованием критерия оптимального ортогонального разложения и оценкой на этой основе разложения обычного.

PACS: 02.30.-Mv; 02.60.-Ed; 02.60.-x

INTRODUCTION

The water spectra method applies a drop taken from a water probe to measure the water's state spectrum. In the special experiment, the drop is placed on a hostaphan folio (Fig. 1). During the whole process of evaporation of the drop, one measures at equal time intervals the drop contact angle with the folio. On independent axis one has the values of the contact angles within fixed angular intervals and on dependent axis — the frequency of measurements of these angles. $\phi(\theta)$ is the distribution of the number of measurements of contact angle θ during the drop evaporation. One can change the function $\phi(\theta)$ of the angle θ to the function of energy variable $F(E)$ using the following Antonov transformation [1, 2]:

$$E(\theta) = \frac{-\gamma(1+\cos\theta)}{I(1+\cos\theta_0)}, \quad F(E) = \frac{b\phi(\theta)}{\sqrt{1-(1+bE)^2}},$$

where $b = I(1+\cos(\theta_0))/\gamma$. Here, $I = 5.03 \cdot 10^{18}$ m^{-2} is the density of water molecules in the surface layer; γ is the surface tension; θ_0 — the initial contact angle. The so-obtained graph after measurements by the method in [1] is referred to as energy spectrum $F(E)$ of the probe. The domain of $F(E)$ contains the values of the hydrogen bond energy.

[1] E-mail: nibogd@inrne.bas.bg
[2] E-mail: todorov_st@yahoo.com

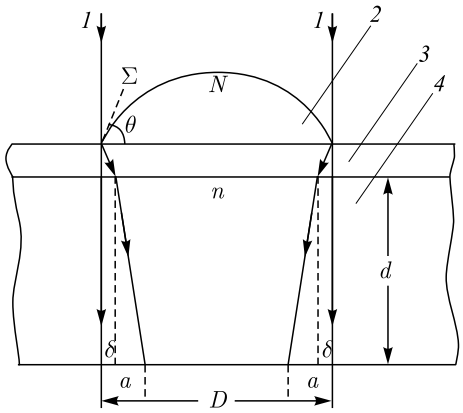

Fig. 1. Experimental setup

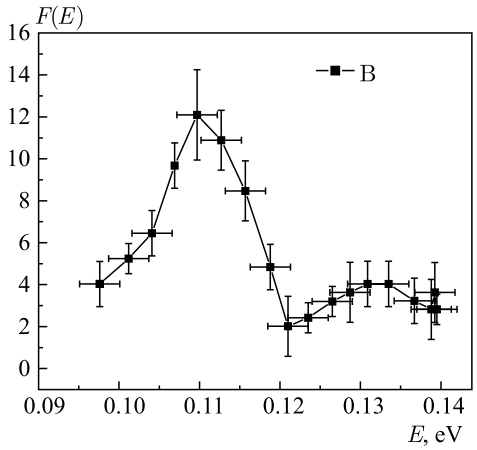

Fig. 2. Given data (B) with errors

The method of water spectra is sensitive to treatment by physical fields as, e.g., γ-ray treatment of water filtration [1, 2] and to environmental changes of the ecosystem and to mechanical treatment and aeration on different water probes [3, 4]. In the present paper, we approximate another natural water data taken from a water spring near the village Lenovo in Bulgaria. In Fig. 2, we present new detailed information about the given data and their errors in both variables of water probe.

1. MAIN PROBLEM DEFINITION

• To find the best approximation curve of the measured water data in Fig. 2, taking into account the errors in both variables.

• To extend our Orthonormal Polynomial Expansion Method (OPEM), according to some criteria, to evaluate orthonormal description of the given data. To find the best approximating curve with usual polynomials, received by the orthonormal ones, according to some criteria.

2. NUMERICAL METHOD — OPEM "TOTAL VARIANCE"

Let $\{E_i, F_i, i = 1, \ldots, M\}$ be the arbitrary pairs of monitoring data E and F, introduced in this section. They are given with experimental errors in both variables — $\sigma(F_i)$ and $\sigma(E_i)$. Consider the total uncertainty (total variance) $S^2(E, F)$ [5–7], associated with (E, F),

$$S_i^2 = \sigma^2(F_i) + \left(\frac{\partial F_i}{\partial E_i}\right)^2 \sigma^2(E_i), \tag{1}$$

according to the ideas of Bevington (1969) [5], where his proposal was to combine the errors in both variables and to assign them to dependent variable. One defines the errors corridor $C(E, F)$, which is the set of all intervals

$$[F(E) - S(E, F), F(E) + S(E, F)]. \tag{2}$$

2.1. Orthonormal Expansion Criteria. The first criterion to be satisfied is that the fitting curve should pass within the errors corridor $C(E, F)$. In the cases of errors only in F (i.e., $\sigma(E) = 0$, $\sigma(F) \neq (0)$), the errors corridor $C(E, F)$ reduces to the known set of intervals

$$[F - \sigma(F), F + \sigma(F)] \tag{3}$$

for any F. The second criterion is that the fitting curve $F^{\mathrm{appr}}(E_i)$ satisfies the expression

$$\chi^2 = \sum_{i=1}^{M} \frac{w_i[F^{\mathrm{appr}}(E_i) - F(E_i)]^2}{M - L}, \quad w_i = \frac{1}{S_i^2} \tag{4}$$

should be minimal (L — number of polynomials). The preference is given to the first criterion. When it is satisfied, the search of the minimal chi-squared stops. Some details of the calculation procedure are given in Forsythe's paper [7] and in our works [8–10].

Our procedure gives results for approximating function by two expansions: of orthogonal coefficients $\{a_i\}$ with optimal degree L_a and of the usual ones $\{c_i\}$ with optimal degree L_c:

$$F^{\mathrm{appr}(m)}(E) = \sum_{i=0}^{L_a} a_i P_i^{(m)}(E) = \sum_{i=0}^{L_c} c_i E^i. \tag{5}$$

The orthogonal coefficients are evaluated by the given values F_i, weights and orthogonal polynomials:

$$a_i = \sum_{k=1}^{M} F_k w_k P_i^{(m)}(E_k). \tag{6}$$

Our recurrence relation for generating orthonormal polynomials and their derivatives ($m = 1, 2, \ldots$) (or their integrals with $m = -1, -2, -3, \ldots$) are carried out by

$$P_{i+1}^{(m)}(E) = \gamma_{i+1}[E - \mu_{i+1})P_i^{(m)}(E) - (1 - \delta_{i0})\nu_i P_{i-1}^{(m)}(E) + mP_i^{(m-1)}E)], \tag{7}$$

where μ_i and ν_i are the recurrence coefficients, and γ_i is the normalizing coefficient, defined by scalar products of the given data. One can generate $P_i^m(E)$ recursively. The polynomials satisfy the following orthogonality relation:

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

over the discrete point set $\{E_i, i = 1, 2, \ldots, M\}$. All the calculations for the sake of uniformity are carried out for E in $[-1, 1]$, i.e., after the input interval is transformed to the unit interval. We remark some advantages of OPEM: it uses unchanged the coefficients of the lower-order polynomials; it avoids the procedure of inversion of the coefficient matrix to obtain the solution. All these features shorten the computing time and assure the optimal solution by the criteria (2) and (4).

2.2. Usual Expansion Criteria. The inherited errors in usual coefficients are given by the inherited errors in orthogonal coefficients:

$$\Delta c_i = \left(\sum_{k=1}^{L} (c_i^{(k)})^2 \right)^{1/2} \Delta a_i, \tag{8}$$

$$\Delta a_i = \left[\sum_{k=1}^{M} P_i^2(E_k) w_k (F_k - F_k^{\mathrm{appr}})^2 \right]^{1/2}, \tag{9}$$

where coefficients $c_i^{(k)}$ are defined by the expansion of polynomials in ordinary polynomial basis

$$P_k = \sum_{i=0}^{k} c_i^{(k)} E^i, \quad k = 0, \ldots, L \tag{10}$$

and explicitly constructed by recurrence relation in [12].

The procedure is iterative because of the evaluation of derivatives on every iteration step. We note that in every iteration step the algorithm finds the best approximation using given before criteria. Having the optimal degree L_a, we continue with finding the optimal L_c. We are asking for the *minimal value* of the maximal distance between functions, evaluated by orthonormal and usual expansions through all iterations,

$$\max |(F_{a,k}^{\mathrm{appr}} - F_{c,k}^{\mathrm{appr}})|, \quad k = 1, \ldots, M. \tag{11}$$

3. APPROXIMATION RESULTS

The main important results with approximation degrees $2{-}10$ for iterations $1{-}9$ are presented in Table 1 for characteristics: number of iterations, number of polynomials, χ^2, and $\max |(F_a - F_c)|$. We see from the table that for iterations $2{-}5$ with optimal number $L_a = 6$, the results are good for both expansions, and for the 8th iteration with optimal number $L_c = 8$, the usual expansion is also good. We present in figures the three curves — given (B), approximated by orthogonal polynomials (C), and received from it by usual polynomials (D) at different iteration steps. Below, Figs. 3 and 4 show the orthogonal (C) and usual (D) approximations for the 4th ($L = 6$) and the 8th ($L = 5$) iterations.

Table 2 presents the approximation at the 5th degree in the 8th iteration of $M = 18$ given values of the following characteristics: energy E, distribution F, σ_E and $\sigma_{F(E)}$, and from the 5th column — approximating values with orthonormal coefficients $F_a^{\mathrm{appr},5}$, approximating values with usual coefficients $F_c^{\mathrm{appr},5}$, differences $\Delta(F_a, F_c) = (F_a^{\mathrm{appr},5} - F_c^{\mathrm{appr},5})$, total

Table 1. **OPEM approximations results for every step approximation**

Characteristics	k^{it}								
	1	2	3	4	5	6	7	8	9
L $(2{-}10)$	7	6	6	6	6	5	6	5	6
$\chi^2 \cdot 10^{-1}$	5.61	4.24	3.99	3.79	3.77	6.81	3.75	6.65	3.63
$\max \lvert(F_a - F_c)\rvert$	14.96	3.48	6.75	4.80	4.63	0.03	6.16	0.08	9.33

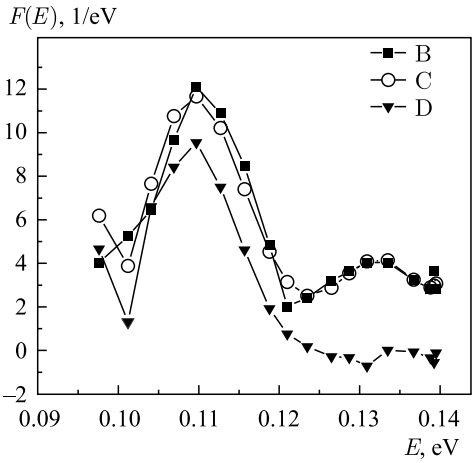

Fig. 3. Orthogonal approximation (C) and the usual one (D) $L = 6$, $k^{\rm it} = 4$

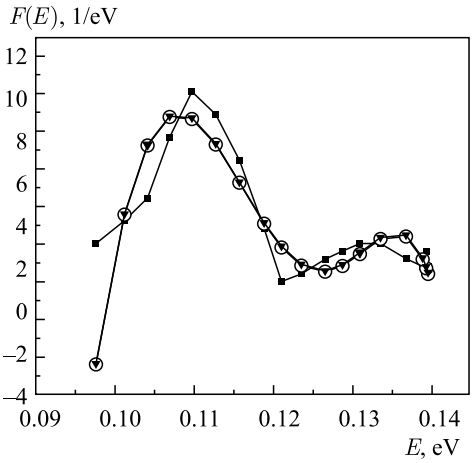

Fig. 4. Orthogonal approximation (C) and the usual one (D) $L = 5$, $k^{\rm it} = 8$

Table 2. **OPEM approximation of contact water energy data**

No.	E, eV	$F(E)$	σ_E	σ_F	$F_a^{\rm appr,5}$	$F_c^{\rm appr,5}$	$\Delta(F_a, F_c)$, $\times 10^{-2}$	S
1	0.1395	2.820	0.025	0.72	2.421	2.503	8.169	2.2072
2	0.1392	3.627	0.025	1.43	2.721	2.799	7.796	2.9469
3	0.1388	2.822	0.025	1.43	3.192	3.266	7.420	2.2173
4	0.1367	3.227	0.025	1.08	4.408	4.484	7.614	1.8114
5	0.1335	4.035	0.025	1.08	4.272	4.353	8.125	1.3297
6	0.1309	4.035	0.025	1.08	3.467	3.549	8.161	1.3126
7	0.1287	3.632	0.025	1.43	2.840	2.905	6.474	2.6050
8	0.1265	3.200	0.025	0.72	2.534	2.583	4.910	0.9395
9	0.1235	2.422	0.025	0.72	2.861	2.932	7.089	0.5500
10	0.1210	2.017	0.025	1.43	3.821	3.889	6.886	3.4402
11	0.1188	4.840	0.025	1.08	5.091	5.137	4.575	5.1487
12	0.1157	8.470	0.025	1.43	7.259	7.291	3.272	8.2753
13	0.1127	10.887	0.025	1.43	9.290	9.334	4.365	5.3774
14	0.1097	12.095	0.025	2.15	10.647	10.700	5.320	4.6238
15	0.1069	9.677	0.025	1.08	10.750	10.793	4.292	6.4789
16	0.1041	6.452	0.025	1.08	9.243	9.276	3.293	15.8508
17	0.1012	5.242	0.025	0.72	5.569	5.601	3.178	6.0766
18	0.0975	4.030	0.025	1.08	–2.384	–2.347	3.714	86.5354

variance $S(5)$ (Eq. (1)). Table 2 shows good coincidence between two descriptions. For comparison we can see the previous results for OPEM applications in [10–12].

CONCLUSIONS

We have developed a new version of OPEM algorithm (7) and Fortran77 package to include errors in both variables defined as new "total variance" (1), according to (2) and (4), taking into account the inherited errors (8), (9) in coefficients.

The results show that the orthonormal and usual expansions values (5) are close to the given ones in the whole interval. The optimal approximations of contact (wetting) energy show good *accuracy and stability*. We received suitable descriptions of the energy variations useful for further investigations with typical formes of approximated curves for different water probes.

REFERENCES

1. *Antonov A., Todorova L.* // Compt. Rend. Acad. Bulg. Sci. 1995. V. 48. P. 21–24.
2. *Todorova L., Antonov A.* / Compt. Rend. Acad. Bulg. Sci. 2000. V. 53. P. 43–45.
3. *Antonov A. et al.* // Commun. Franco-Bulgare OM. 1993. V. 1.
4. *Bonn D., Ross D.* Wetting Transitions // Rep. Prog. Phys. 2000. V. 64. P. 1085;
   Fuchs N. A. Evaporation and Droplet Growth in Gaseous Media. London: Pergamon, 1959.
5. *Bevington P. R.* Data Reduction and Error Analysis for the Physical Sciences. New York: McGrow-Hill, 1969.
6. *Jones G.* Preprint TRI-PP-92-31, A. 1992;
   Orear J. // Am. J. Phys. 1982. V. 50. P. 912;
   Lybanon M. // Am. J. Phys. 1984. V. 52. P. 276.
7. *Forsythe G.* // J. Soc. Ind. Appl. Math. 1957. V. 5. P. 74–87.
8. *Bogdanova N.* JINR Commun. E11-98-3. Dubna, 1998.
9. *Bogdanova N., Todorov St.* // IJMPC. 2001. V. 12. P. 117–127.
10. *Bogdanova N.* // BPU6 Conf., Istanbul, Aug. 2006. 2007 AIP Proc. / Ed. by S. A. Cetin, I. Hikmet.
11. *Bogdanova N., Todorov St.* // BPU7 Conf. Alexandroupolis, Greece, Sept. 2009. 2010 AIP Proc. / Ed. by A. Angelopoulis, F. Takis.
12. *Bogdanova N., Todorov St.* // MMCP 2009, Dubna, LIT. Bull. PFUR. Ser. Math. Inform. Sci. Phys. 2011. V. 3, No. 2. P. 63–67.