ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

Дубна

E7-99-273

V.B.Zlokazov

MATHEMATICAL ANALYSIS OF DATA IN THE EXPERIMENT ON THE SYNTHESIS OF THE ELEMENT 114

Submitted to «European Physical Journal A»

1999

Злоказов В.Б. Е7-99-273
Математический анализ данных эксперимента
по синтезу 114 элемента

Проведен статистический анализ зарегистрированной цепочки сигналов имплантации ядра отдачи, трех последовательных альфа-распадов и заключительного спонтанного деления. С помощью критериев статистики оценено правдоподобие следующих гипотез о сигналах: 1) данные — продукты распада элемента 114; 2) данные — случайная комбинация сигналов; 3) данные — продукты распада элемента 112; 4) данные — продукты реакций переноса между ядрами пучка и мишени.

Показано, что значимым образом данные не противоречат лишь первой гипотезе.

Работа выполнена в Лаборатории ядерных реакций им. Г.Н.Флерова ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна, 1999

Zlokazov V.B. E7-99-273
Mathematical Analysis of Data in the Experiment
on the Synthesis of the Element 114

The paper describes the statistical analysis of a registered chain of signals for an implantation of a nucleus recoil, the three consequent alpha decays and the terminating spontaneous fission. With the help of statistical tests the likelihood is estimated of the following hypotheses: 1) data is product of the decay of the element 114; 2) data is a random combination of signals; 3) data is product of the decay of the element 112; 4) data is products of the transfer ractions between the nuclei of the beam and the target.

It is shown that the data doesn't contradict significantly only to the first hypothesis.

The investigation has been performed at the Flerov Laboratory of Nuclear Reactions, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna, 1999

MATHEMATICAL ANALYSIS OF DATA IN THE EXPERIMENT ON THE SYNTHESIS OF THE ELEMENT 114

The logic and the apparatus of nuclear experiments getting ever more complicated, the situation arises when result of such experiments is the observation of one single event, which admits a multiple interpretation, and first of all, as a random signal combination. The direct use of statistical methods in this case is either impossible, or inappropriate: a combination with methods of probability theory is needed. Of course, the mathematical analysis in such situation looses the reliability and safety of the classical statistics; but it allows us to extract the optimum volume of information from the data of a small size which is possible in this case at all[1].

As example, let us consider the analysis of data obtained in the experiment on the synthesis of superheavy nuclei in the $^{48}Ca +^{244} Pu$ reaction [2], in particular, of the 114th element.

The data of interest (in a group of others, which are uninteresting) is a chain of registered signals, starting with the implantation of the recoil nucleus, followed by 3 alpha - decays, and ending with the spontaneous fission; all the signals were observed in the same strip of the detector.

Of course, the decision about the physical nature of the events can be made only by the pysicist; the role of mathematics on condition of exclusively poor statistics consists in testing the correspondence of certain fragments of the quantitative analysis of these data by physicists to criteria of probability theory and mathematical statistics.

Below some mathematical tools needed for the further analysis are described.

Functions of probability distribution for the radioactive decay.

The classical function of probability distribution for an event (the radioactive decay) at a time moment t is: $P(t) = 1 - exp(-lt)$, where $l = ln(2)/T$, and $T =$ the halflife of the nucleus. The density of this probability is $f(t) = l \cdot exp(-lt)$. With the help of $P(t)$ and $f(t)$ all the other distributions can be obtained.

So the function of the probability distribution for the daugther nucleus is

$$P_{12}(t) = \int_0^t f_1(\tau)P_2(t-\tau)d\tau$$

and substituting the concrete expressions in $P_2(t)$ and $f_1(t)$ we get

$$P_{12}(t) = \begin{cases} 1 - exp(-l_1t) - l_1/(l_1 - l_2)(exp(-l_2t) - exp(-l_1t)), & \text{if } l_2 \neq l_1; \\ 1 - exp(-l_1t) - l_1t \cdot exp(-l_2t); & \text{if } l_2 = l_1 \end{cases}$$

Similarly, the same functions are built for the successors of the consequent decay of the original nucleus:

$$P_{123}(t) = \int_0^t f_1(\tau)P_{23}(t-\tau)d\tau, \quad (1)$$

$$P_{1234}(t) = \int_0^t f_{12}(\tau)P_{34}(t-\tau)d\tau, \quad (2)$$

$$P_{12345}(t) = \int_0^t f_{123}(\tau)P_{45}(t-\tau)d\tau, \quad (3)$$

where $1, 2, 3, 4, 5$ denote mother, daugther, grand daugther etc., respectively. The concrete formulae for (1, 2, 3) are too bulky due to many combinations of coinciding and not coinciding halflives and here are omitted.

Function of probability distribution for a quadratic form.

Let the following quadratic form

$$S_n = \sum_{i=1}^{n} x_i^2, \tag{4}$$

be given, where normally distributed random quantities x_i have zero expectation, unit variance, and the neighbouring pairs x_i,x_{i+1} are correlated with the correlation coefficient -0.5. What is the function of probability distribution of S_n? One of the widely spread errors in the practice of data analysis is that this function is supposed to have the χ_n^2 distribution. In fact, it has not.

From the mathematical statistics [4] it is known that the quadratic form $Q_{ij}x_ix_j$, $i,j = 1,...,n$, where Q is the inverse normal covariance matrix of x_i, has the χ_n^2-distribution, irrespective of whether the x_i are correlated or not; therefore, if we had taken the inverse matrix to

$$c_{ij} = \begin{pmatrix} 1 & -0.5 & 0 & ... & 0 \\ -0.5 & 1 & -0.5 & ... & 0 \\ & & ... & & \\ 0 & 0 & ... & 1 & -0.5 \\ 0 & 0 & ... & -0.5 & 1 \end{pmatrix}. \tag{5}$$

and built a quadratic form

$$\hat{S}_n = \sum_{i,j=1}^{n} c_{ij}^{-1} x_i x_j,$$

it would have the χ_n^2-distribution. But the inversion of a matrix like (5) is rather complicated; besides, the probability of large deviations of a χ^2-distributed random quantity is substantially larger than that of S_n (e.g., for $n = 4$ about 0.07 and 0.045, respectively) and the decision making procedure based on the use of $\hat{S}_n$ loses part of its efficiency specifically in this case.

Thus, preferable is a method based on the direct use of S_n. The probability distribution function and its density for (4) can be easily calculated numerically. One can show, that the expectation of (4) is equal to n, and the variance to $3n - 1$. Below a table of values of $P(t)$ for $n = 4$ is given.

P	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
S_4	.6	.9	1.1	1.4	1.6	1.9	2.2	2.5	2.8	3.1
P	.55	.60	.65	.70	.75	.80	.85	.90	.95	
S_4	3.4	3.8	4.2	4.7	5.3	6.0	6.9	8.2	10.3	

Numbers in each upper row are probabilities P with the step 0.05, and in the lower one - the corresponding values of S_4.

With the help of such table (certainly, more detailed) we can construct the 67%- confidence interval of S_4 as $n \pm \sigma$; corrected for the asymmetry it is: (1.8 - 10.8).

The formalism of stochastic Poisson time processes.

These are the time functions $K(t_1, t_2)$ - number of random events, occured during a time interval (t_1, t_2) with a probability $Q_k(t_1, t_2)$ and having the following properties:

1. stationarity: $Q_k(t_1, t_2) = Q_k(t_2 - t_1)$ for arbitrary t_1, t_2;
2. $Q_k(t_1, t_2)$ independence of the event prehistory: $Q_k(t_1, t_2|C) = Q_k(t_1, t_2)$, where C means events which happened before t_1;
3. rareness of events: $Q_{k>1}(\delta t) = o(\delta t)$.

These properties allow us to write simply $K(t)$ and $Q_k(t)$ bearing in mind that t means the duration of the time interval considered.

The Poisson processes play an important role in analytical modelling of the stochastic background in scientific and technical applications because random events very often satisfy the above-numbered requirements. They disable the need to use the computer simulation to get the estimates of the random background characteristics.

The function of probability distribution of $K(t)$ is

$$Q_k(t) = \frac{(lt)^k}{k!} exp(-lt), \tag{6}$$

where l = parameter of the Poisson distribution, t = time, and Q_k = probability, that during a time interval $(0, t)$ k events will be registered.

The quantity lt is the expectation and at the same time the variance of $K(t)$ at a moment t.

DATA ANALYSIS

Now we start the analysis of the data chain mentioned above. We shall go over a set of possible interpretations of this data proposed by the physicists and consider the following problems: within the framework of these interpretations estimate the formal probabilities of the observed signal configuration and some of its statistical characteristics.

Interpretation 1: "data is the result of the decay of element-114 recoil".

If this interpretation is valid the data is a sequence of the events: implantation of the recoil, 3 consequent alpha-decays and finally the spontaneous fission.

The quantitative analysis consists in the following:

1. test the correspondence of the observed energies and halflives of the alpha - particles to calculations, given, e.g., in [3].
2. estimate the probabilities of the observed chain of signals for different interpretations and compare them.
3. test the hypothesis about the genetic connection of the signals.

The problem 1 can not be solved by the formal methods because of the absence of reliable information about the required statistical distributions. One can make only statements of qualitative character, e.g., that the observed energy values (from the range 8-10 Mev) and halflives (from the range of several minutes) correspond to the calculations [3]. Therefore, we will focus on the problems 2 and 3.

Unfortunately, we don't know the halflives of the nuclei produced in the decay, but we have a priori estimates of the intervals containing these halflives, and we can set the problem as follows: determine the maximum and minimum probability of the decay of grand grand daugther P_{1234} (2) within the time range from the signal of recoil implantation to the signal of the spontaneous fission over the direct product of confidence intervals for the halflives.

The calculation of the maximum and minimum of (2) over the region

$$0.05 \le T_1 \le 0.5; \quad 30 \le T_2 \le 300; \quad 2 \le T_3 \le 20; \quad 7 \le T_4 \le 27;$$

in the time interval (0, 34min) with account of the registration efficiency, equal to 0.87, gave the following results:

$$P_{min} = 0.0083, \quad P_{max} = 0.3364.$$

The problem 3 can be solved in the following way. For testing a hypothesis about the genetic connection between the signals we have the following data about the locations of the decaying nucleus in the position - sensitive detector:
xevr = 16.5mm - position of the signal: implantation.
xa1 = 15.6mm - position of the signal: alpha -1.
xa2 = 16.5mm - position of the signal: alpha -2.
xa3 = 17.0mm - position of the signal: alpha -3.
xsf = 17.1mm - position of the signal: spontaneous fission.

Then we have "resolutions" - FWHMs of the distributions of signal differences, which, on assumption of the normal distribution of these differences, can be transformed to the usual sigmas:

EVR - alpha: resolution = 1.4 mm; $sig_{evr-a} = 0.59$ mm;
alpha - alpha: resolution = 1.0 mm; $sig_{a-a} = 0.42$ mm;
EVR - SF: resolution = 1.2 mm; $sig_{evr-sf} = 0.51$ mm;

Let a hypothesis be tested: all the signals arise as a result of a decay of a parent nucleus, which is located at a fixed position in the detector strip.
The statistical test of this hypothesis can be carried out by two methods.
Method 1. Let us construct an expression

$$S = (\frac{xevr - xa1}{sig_{evr-a}})^2 + (\frac{xa1 - xa2}{sig_{a-a}})^2 + (\frac{xa2 - xa3}{sig_{a-a}})^2 + (\frac{xevr - xsf}{sig_{evr-sf}})^2 \qquad (7)$$

Substituting the corresponding values of variables into (7), we get: $S_4 = 9.56$. One sees at once, that it is covered by the 67% - confidence interval of the quantity S_4: (1.8 - 10.8).
Method 2. The analysis of the differences is less efficient than the analysis of their constituents, since the variance of the formers is always greater than that of the latters. Besides, these variances are obtained from the calibration reactions, and can differ from the true variances of difference signals for the reaction considered.
Therefore, for a greater reliability we can use the classical approach of the statistics: analysis of the constituents of these differences. Let us find the sample mean and the variance for the signals of the nucleus position. We have for the mean:

$$pos = (16.5 + 15.6 + 16.5 + 17.0 + 17.1)/5 = 16.54;$$

For the sample variance we apply the usual formula:

$$var = \sum_{i=1}^{n}(x_i - pos)^2/(n-1)$$

After the necessary calculations we obtain $var = 0.35$, whence we find the sigma: $\sigma = 0.59$.
Let us consider the expression

$$Q = \sum_{i=1}^{5}(\frac{x_i}{\sigma})^2 \qquad (8)$$

where x_i = difference between ith signal and pos.

If our hypothesis holds then under very common assumptions each quotient and their sum will have asymptotically the Student's and χ^2_3 distributions, respectively. Substituting our data in Q, we get $Q = 4.0$.
The expectation of this quantity is $\hat{E}Q = m = 3$, the variance $\hat{V}Q = 6$ and $\sigma = 2.45$. The 67%- confidence interval calculated as $m \pm \sigma$ and corrected for the asymmetry is equal to (1.30, 7.50); our Q gets into it. Thus, both methods find out that the data does not contradict to the above hypothesis.
REMARK. The variance of the difference is twice greater than the variance of its constituents; since the sigmas of the differences are smaller than 0.59, it points out that real difference variances are larger than the given above. So, the real value of (7) is even smaller than 9.56.
A special question: why $x_2 = (xa1 - pos) = -0.94$ is 1.6 times greater than σ, equal to 0.59?
The probability that a Student 4 degrees of freedom distributed quantity gets into the interval $(1.6 \cdot \sigma, \infty)$ is between 5% and 10% - such probability is quite admissible for the statistical fluctuations. Anyway, to offer another hypothesis about the alpha-1 signal emergence additional evidence is needed - the statistics of the above - considered data is not enough.

Interpretation 2: " events are random, and the spontaneous fission has no relation to the reaction $^{48}Ca +^{244} Pu$".
The calibration measurement of chance signals of recoil implantation and alpha - particles with energy 8.5-10 in a detector strip for a position-correlation window 1.6mm gave the following frequencies:
implantation = 1,3 per hour, alpha - particle = 1 per hour.
From this data we can derive the probability of the events: one imitator of the implantation signal 34 minutes before the spontaneous fission, and 3 imitators of alpha - particles between them.
The above-mentioned random events represent a typical time process of the Poisson type(6). We can estimate l in (6) for implantation and and alpha - particle imitators on the basis of calibration data as follows:

$$l_i \cdot 60 = 1.3; \quad l_a \cdot 60 = 1.$$

Solving these equations we get $l_i = \frac{1,3}{60}$; $l_a = \frac{1}{60}$.
Substituting in (6) we get the probability of 3 alpha - particles

$$Q_3(34) = \frac{(34/60)^3}{3!} exp(-(34/60)),$$

and the probability of one implantation 34 minutes before the spontaneous fission:

$$P_r(34) = (34 \cdot 1.3/60) exp(-(34 \cdot 1.3/60))$$

Thus, we have the probability P_s for this data interpretation

$$P_s = Q_3 \cdot P_r \sim 0.00607$$

This probability does not give yet a notion about the likelihood of the considered signal interpretation - to make a statistically correct decision it is necessary to compare it with the probabilities of other random signal combinations. We have:

Imp.	Alpha	Probability
1	0	0.20010
1	1	0.11339
1	2	0.03213
1	4	0.00086
2	1	0.04176
2	0	0.07370

Here Imp = number of signals for the implantation , Alpha = for the number of alpha - particles, Probability = probability of such combination. It is seen that the largest is the probability to observe the combinations 1+0 and 1+1, but the probability of the combination 1+3 is really small as compared with them. In other words, the data obviously contradicts to the hypothesis about the random character of signal emergence.

Interpretation 3: "the chain is decay of the element 112, and one alpha - particle is imitator".
Assuming that the recoil nucleus is the element 112 - by $(\alpha, 3n)$-evaporation channel - and the first alpha - particle is imitator, the problem is: determine the maximum and minimum probability of the decay of the grand daughter P in the time interval from implantation signal to that of the spontaneous fission on the direct product of confidence intervals for the halflives.
The maximum and minimum of (1) over the region

$$30 \leq T_2 \leq 300; \quad 2 \leq T_3 \leq 20; \quad 7 \leq T_4 \leq 27;$$

in the time interval (0, 34min) gave the following results:

$$\hat{P}_{min} = 0.0293, \quad \hat{P}_{max} = 0.4934.$$

Multiplying these probability by the probability of the imitation of one alpha - particle in the time interval (0, 34min) $Q_1(34) = \frac{34}{60} exp(-(34/60))$, and correcting them for the registration efficiency, we finally get

$$P_{min} = 0.0032, \quad P_{max} = 0.1273.$$

Interpretation 4: "the chain is the decay of a product of the transfer reaction between the nuclei of the projectile and the target, and alpha - particles (all or part) are imitators".
This case is similar to the previous one, but the probabilities of the event configurations and the genetic connection between them will be smaller, and the more alpha - particles are suppposed to be imitators, the smaller. As instance, let us consider the case: 1 alpha - particle is true, the other 2 are imitators. Suppose that the halflife of the mother - nucleus is contained in the interval (1 - 100min). Omitting the details of the calculations (they are similar to the above ones), we get

$$P_{min} = 0.0089, \quad P_{max} = 0.1521.$$

The analysis of the results. We have:

Interpretation number	Max. probability	Min. probability
1	0.3364	0.0083
3	0.1273	0.0032
4	0.1521	0.0083

One can see that the comparison of the formal probabilities to observe the signal configuration given does not contradict to the preference for the interpretation 1 made by the authors of [2]. And this non-contradiction substantially increases if we attach the physical probabilities, by which the formal probabilities should be multiplied:

1. The cross-section of the channel $(\alpha, 3n)$ for interpretation 3 is several orders smaller than that of the channel $(3n)$ for interpretation 1;
2. Probabilities to observe the given in [2] energies of alpha - particles and possible halflives for interpretation 4 are very small as compared with the probabilities for interpretation 1.

Unfortunately, the numerical evaluation of the physical probabilities is impossible, since the calculations like [4] don't contain confidence intervals for the possible energies of the alpha - decay and halflives.

Conclusion

The performed analysis shows the importance of developping the following aspects of the mathematical data treatment in case of rare events:
1. the formalism for the analytical modelling of the stochasic background;
2. the mathematically grounded calculus of hypothesis probabilities.

Acknowledgements.

This work has been performed at the request of Prof. Yu.Ts.Oganessian. The author thanks him and Drs. V.K.Utyonkov and A.G.Demin for the helpful discussions of the physical aspects of the problem.

References.

1. K.H.Schmidt, C.C.Sahm, K.Pielenz, H.G.Clerc. Some Remarks on the Error Analysis in the Case of Poor Statistics. Z.Phys.A - Atoms and Nuclei 316, 19-26 (1984).

2. Yu.Ts.Oganessian, V.K.Utyonkov, et al. The Synthesis of Superheavy Nuclei in the $^{48}Ca + ^{244}Pu$ Reaction. JINR Preprint E7-99-53, Dubna, 1999

3. R.Smolanczuk. Properties of the hypothetical spherical superheavy nuclei. Phys.Rev. C 56 N2, p.812 (1997).

4. Wilks S.S. Mathematical Statistics, N.Y.-Lond; J.Wiley, 1962.

Received by Publishing Department
on October 18, 1999.