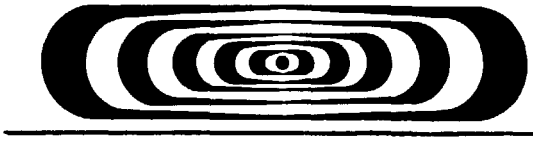
ISSN 0101-3084



CNEN/SP

ipen Instituto de Pesquisas Energéticas e Nucleares

LEAST SQUARE METHODS AND COVARIANCE MATRIX APPLIED TO THE RELATIVE EFFICIENCY CALIBRATION OF A Ge(Li) DETECTOR

Luiz Paulo Geraldo and Donald L. Smith

PUBLICAÇÃO IPEN 243 JANEIRO/1989

SÃO PAULO

PUBLICAÇÃO IPEN 243 JANEIRO/1989

LEAST SQUARE METHODS AND COVARIANCE MATRIX APPLIED TO THE RELATIVE EFFICIENCY CALIBRATION OF A Ge(Li) DETECTOR

Luiz Paulo Geraldo and Donald L. Smith

DEPARTAMENTO DE FÍSICA E QUÍMICA NUCLEARES

CNEN/SP
INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES
SÃO PAULO – BRASIL

Série PUBLICAÇÃO IPEN

INIS Categories and Descriptors

F50.00

E41.40

CALIBRATION
LEAST SQUARE FIT
LI-DRIFTED GE DETECTORS
MATRICES

IPEN - Doc - 3121

Aprovado para publicação em 26/12/88.

LEAST SQUARE METHODS AND COVARIANCE MATRIX APPLIED TO THE RELATIVE EFFICIENCY CALIBRATION OF A Ge(Li) DETECTOR

Luiz Paulo Geraldo and Donald L. Smith *

ABSTRACT

The methodology of covariance matrix and least square methods have been applied in the relative efficiency calibration for a Ge(Li) detector. Procedures employed to generate, manipulate and test covariance matrices which serve to properly represent uncertainties of experimental data are discussed. Calibration data fitting using least square methods has been performed for a particular experimental data set.

MÉTODOS DE MÍNIMOS QUADRADOS E MATRIZ DE COVARIÂNCIA APLICADOS NA CALIBRAÇÃO DE EFICIÊNCIA RELATIVA PARA UM DETECTOR DE Ge(Li)

RESUMO

A metodologia de matriz de covariância e métodos de mínimos quadrados foram aplicados na calibração de eficiência relativa para um detector de Ge(Li). Os procedimentos empregados para gerar, manipular e testar matrizes de covariâncias, que servem para representar apropriadamente incertezas de dados experimentais, são discutidos. O ajuste dos dados de calibração foi efetuado, para um particular conjunto de resultados experimentais, utilizando métodos de mínimos quadrados.

INTRODUCTION

The analysis of experimental nuclear data and the evaluation of the data base for applications are matter which have been receiving a lot of attention recently. The incresingly severe data accuracy requirements have led researchers to examine carefully the methods used in processing nuclear data. The main objectives are the elimination of

(*) Argonne National Laboratory - Engineering Physics Division Building, 314 - Argonne, Illinois 60439, U.S.A.

bias in evaluation, and realistic assessment of uncertainties in the data base. The mathematical procedures to handle this problem have been available since the time of Gauss (1809). However, only recently that a more rigorous treatment of experimental data involving covariance matrix and least square methods has been introduced by Perey (1)-(2). More detailed studies on this subject addressing to neutron nuclear data have been presented in the literature (3)-(6). These studies have shown a tendency to be extended for other nuclear data fields as well as for other experimental researches (7).

The main purpose of this paper is to acquaint readers of any area with a few basic ideas and to offer some references which provide the opportunity to learn more about this important subject. These goals will be pursued by presenting the methodology for a real and simple example of relative efficiency calibration for a Ge(Li) detector, a procedure which has shown to be important for example, in cross section measurements using the activation technique. The methodology description is presented in a sufficiently self contained form to enable the reader to follow, in a logical sequence, the essential steps for the application of this methodology, without having to refer constantly to the related papers.

FORMALISM

One important step to be concerned in any experiment today is the development of a covariance matrix for the set of experimental parameters which are deduced from measurements in the laboratory. The covariance matrix is a more complete form of uncertainty representation than the older statistic methods because besides the total error it gives information about the existing level of correlation among the parameter errors. Due to the latter fact, very often the covariance matrix is presented in the literature in the following equivalent form: the total errors and the correlation matrix, separetely.

The method of deriving a covariance matrix for any experimental data has been discussed in some detail by SMITH (3), so it will not be pursued here in any depth. Summarizing briefly, we assume that there

are L distinct sources of error determining the overall uncertainties for a collection of n quantities $\bar{X} = (x_1, x_2, \ldots, x_n)$. Let E_{i1} represent the particular error in x_i corresponding to the attribute 1, and let $\bar{C}_1$ be the correlation matrix corresponding to these partial errors, where C_{ij1} is a typical element, as they apply to all componentes of $\bar{X}$. The coefficient C_{1ij} determines the correlation degree between the errors E_{xi} and E_{xj} due to effect 1 and it is usually defined as micro correlation. The possible values for the elements C_{ij1} must fall within the closed interval (-1,1) with the following significance, according to the value assumed by C_{ij1}: = 0, no correlation; = +1, complete correlation; = -1, complete anticorrelation and, for the intermediary cases partial correlations or partial anticorrelations depending if $0 < C_{ij1} < +1$ or $-1 < C_{ij1} < 0$ respectively. The total correlation matrix representing this set of quantities is formed of elements V_{ij} calculated using the expression:

$$1) \quad V_{xij} = \sum_{\ell=1}^{L} E_{i1} E_{j1} C_{ij1}$$

The final information for the error vector $\bar{E}_x$ and for the total correlation matrix $\bar{C}_x$ can be obtained by the following relationships:

$$2) \quad E_{xi} = (V_{xii})^{1/2} \text{ and } C_{xij} = V_{xij}/(E_{xi} E_{xj})$$

Another point to be remembered is that $\bar{V}_x$ and $\bar{C}_x$ are both symmetric, i.e., $V_{xij} = V_{xji}$ and $C_{xij} = C_{xji}$.

The first step in deriving a covariance matrix is to catalogue all the sources of error existing in the variables x_i and decide what type of correlation exist. These errors should be expressed in units of the corresponding variables so that the result of this exercise would be a reference table as presented in Table 1.

Table 1

variables	error componentes (l)					total error
(i)	1	2	3	.....	L	(E_{xi})
1	E_{11}	E_{12}	E_{13}	.....	E_{1L}	E_{x1}
2	E_{21}	E_{22}	E_{23}	.....	E_{2L}	E_{x2}
"	"	"	"	.....	"	"
"	"	"	"	.....	"	"
n	E_{n1}	E_{n2}	E_{n3}	.....	E_{nL}	E_{nx}

It is evident that some E_{il} may be zero. The partial correlation coefficients C_{ijl} in general are either -1 (rarely), 0 or +1. Intermediary correlations or anticorrelations at this stage are relatively unusual and they represent the most difficult task for the experimenter to decide about them because, it is often necessary to use some subjectivity. After this study is completed it is straightforward to use Eq. 1 to generate the correspondent covariance matrix. However, if we are working with a large number of attributes and/or experimental data points, it is evident that covariance analysis represents a lot of work. In realistic applications, it is often impractical to perform the analysis by hand and on must resort to the use of computers. A FORTRAN computer program CALCOV has been developed to generate a covariance matrix and to produce in addition the correlation matrix and the total errors. It has been implemented on a IBM personal computer (or compatible) and details on the computational procedures can be seen in the Appendix A.

Suppose now that it is intended to know the uncertainty in a quantity F which is a scalar function of the n random variables x_i, denoted collectively as previously by $\bar{X}$. There are several entirely equivalent ways we can express the uncertainty in the quantity F. They are quadratic forms or error propagation formulas expressed by:

$$3)\quad E_f^2 = (\bar{S}\,\bar{E})^t\,\bar{C}_x\,(\bar{S}\,\bar{E}) = (\bar{S}\,\bar{I})^t\,\bar{V}_x\,(\bar{S}\,\bar{I}) = \bar{T}^t\,\bar{V}_x\,\bar{T}$$

The matrix multiplication is assumed in this paper for all expressions involving matrices. The symbol "t" denotes matrix transposition and, the other parameters not defined previously are:

$\bar{I}$ = is a vector with all n elements equal to 1.

$\bar{T}$ = is a vector with typical elements equal to $\partial F/\partial x_i$, and

$\bar{S}$ = is a n x n diagonal matrix where the elements are equal to $\partial F/\partial x_i$.

The information content of matrices $\bar{S}$ and $\bar{T}$ is identical and they are commonly referred to as sensitivity matrices. It is intructive to verify by Eq. 3 that if there is no correlation for all the errors of the variables x_i, then the usual error propagation expression is obtained. For instance, suppose that the function F is given, for a particular case,as

$F = ax_1 + bx_2$ then according to Eq. 3

$$E_f^2 = \bar{T}^t\,V_x\,\bar{T} = \begin{bmatrix}(\partial F/\partial x_1 & \partial F/\partial x_2)\end{bmatrix} \begin{bmatrix}E_{x1}^2 & \\ 0 & E_{x2}^2\end{bmatrix} \begin{bmatrix}\partial F/\partial x_1 \\ \partial F/\partial x_2\end{bmatrix}$$

simple matrix algebra leads to the explicit algebraic expression $E_f^2 = a^2 E_{x1}^2 + b^2 E_{x2}^2$ as predicted by the usual error propagation.

Eq. 3 is a particular case of a broad category of problems concerning covariance matrix in the transformation of variables, which merits some additional discussion. Assume now that one wishes to transform the set of variables x_i to another set of variables f_I and to obtain the appropriate covariance matrix $\bar{V}_f$ for this new set. The transformation is obtained using the matrix $\bar{T}$, also named transformation matrix, in a similar form as discussed above for a scalar function, through the following expression:

$$4) \quad \bar{V}_f = \bar{T}^t \bar{V}_x \bar{T}$$

If the dimension of $\bar{V}_f$ and $\bar{V}_x$ are (n,n) and (m,m) respectively, the matrix $\bar{T}$ has dimension (n,m). The elements of $\bar{T}$ are partial derivatives and given according to the expression:

$$5) \quad T_{kj} = (\partial F_k / \partial x_j) \quad (k = 1,m \text{ and } j = 1,n)$$

If F is a scalar function of the variables x_i then $\bar{V}_f = E_f^2$ as predicted by Eq. 3.

Another approach concerning the previous discussion has been performed using the vector model for error propagation and it is presented elsewhere (8). It yields the same results as the matrix formalism, however it offers a convenient interpretation of the nature of error propagation as well as the significance of the correlation coefficients.

It is essential to comment at this point some physical considerations about the covariance or correlation matrix. The methodology of covariance matrix was introduced when a mathematical model based on statistic methods was employed in the area of experimental data uncertainty. For this reason it is important to ensure that this mathematical model is in conformity with the physical reality (9). It is intuitively clear that all sort of experimental information which one deals in practical situations involve at least some random error and very likely sources of systematic errors as well. Realistic covariance matrices should reflect this situation and therefore they must be positive definite in order to represent uncertainties which are positive and not negative, zero or imaginary. This subject has been treated with some detail in a recent report (10). In summary, the quadratic form represented by Eq. 3 must be positive in order to satisfy the physical reality. On the other hand, it is stated on most text books on matrices that Eq. 3 is positive, for every non-trivial (non-zero) vector $\bar{T}$ if and only if the matrix $\bar{V}_x$ is positive definite. Furthermore, according to these text books, a real symmetric matrix of order n is positive definite if and only if:

i) it has rank equal to the dimension n and all its eigen-values are positive. Eigen-values (λ) of a matrix represent all the solutions to the equation $\det (\bar{V}_x - \lambda \bar{U}) = 0$, where $\bar{U}$ is the identity or unit matrix of dimension n; or

ii) it has rank equal to dimension n and all its leading principal minors are positive. A leading principal minor of a matrix is the determinant of the sub matrix formed by deleting certain rows and the numbered columns.

Using the definitions cited above, two codes MATXTST and MATXTST1 have been developed (10) in order to test a covariance or correlation matrix for positive definiteness property. If the matrix is classified as non positive definite, the programs still provide useful informations concerning the origin of the inconsistency. Covariance matrices generated according to the previous discussion tend to be positive definite so long as the partial errors and their assumed correlations are physically consistent. However, it is a good practice always to test the covariance matrix for this important property in order to avoid possible mistakes or inconsistencies.

It is usual in any experiment that the final experimental information be obtained after the data fitting with some appropriate technique. The most common of these techniques is the least-squares method which will be next discussed, with the formalism of covariance matrices for solving overdetermined systems of linear equations. It will be seen that the covariance methodology can also be very useful in routine data analysis applications. Suppose the following system of n linear equations with n unknown variables:

$$
\begin{aligned}
6)\quad Z_1 &= A_{11} X_1 + A_{12} X_2 + \ldots + A_{1n} X_n \\
Z_I &= A_{I1} X_1 + A_{I2} X_2 + \ldots + A_{In} X_n \\
Z_n &= A_{n1} X_1 + A_{n2} X_2 + \ldots + A_{nn} X_n
\end{aligned}
$$

It may be written in vector notation as:

$$7)\quad \bar{Z} = \bar{A}\,\bar{X}$$

Assuming that $\bar{A}$ is non singular then Eq. 7 has a unique solution which is given by:

8) $\bar{X} = \bar{A}^{-1} \bar{Z}$

In practical situations one is often involved in obtaining the set of best values, which satisfies Eq. 8, from available experimental information $\bar{Z}$. In other words, we are looking for the best solution $X(x_1, \ldots, x_n)$, in the least squares sense, to the following approximate equation:

9) $\bar{Z} \approx \bar{A} \bar{X}$

The symbol "≈" takes into account the approximate relation existing between the experimental information $\bar{Z}$ and the solution $\bar{X}$. Assume that there is a covariance matrix $\bar{V}_z$ which provides the errors and their correlations for $\bar{Z}$. According to the least-squares method, the best possible solution $\bar{X}$ is the one which minimizes the chi-square χ^2. The χ^2 value for this generalized problem is given by:

10) $\chi^2 = (\bar{Z} - \bar{A} \bar{X})^t \bar{V}_z^{-1} (\bar{Z} - \bar{A} \bar{X}) \geq 0$

It is possible to show through caulculus involving matrices (3) that, the condition for obtaining a minimum χ^2 leads to the solution for $\bar{X}$. The desired least-squares solution for $\bar{X}$ is given by:

11) $\bar{X} = \bar{C} \bar{A}^t \bar{V}_z^{-1} \bar{Z}$ where $\bar{C} = (\bar{A}^t \bar{V}_z^{-1} \bar{A})^{-1}$ is the

convariance matrix for the solution $\bar{X}$.

Another important point to be considered is that χ^2 should be distributed in conformity with the chi-square tables, taking into consideration the actual number of freedom degrees f for the problem. This number f is given, in the present case, by the difference (n-m) where n is the number of data input and m is the number of fitted parameters. If χ^2 normalized, i.e., χ^2/f is lower or equal to one then one might assume

that the scatter of the input values is consistent with the assigned errors. On the other hand, a $\chi^2/f > 1$ would represent an inconsistency between the actual scatter of input values and the errors assumed for them. In this case, according to PEELE (11), one possible approach is to perform an adjustment to the solution covariance matrix $\bar{C}$, by multiplying it by χ^2/f or equivalently, by multiplying the correspondent errors by the square-root of this value. This is a crude adjustment and it is justified if the evaluator feels that the input errors were underestimated.

EXPERIMENTAL PROCEDURES

We turn next to an example which will demonstrate (see next section) in the simple terms all of the concepts developed above. This example represents an experiment to obtain the relative efficiency for a Ge(Li) detector, using standard sources produced by different laboratories.

The geometry of the experimental apparatus is shown in Fig. 1 and Table 2 presents the standard sources and respective gamma-rays used for this calibration. For Eu-152 only the gamma-ray of higher energies were considered because, it was of interest to determine the relative efficiencies for gamma-rays of ~ 1 MeV, emmited from some nuclei produced by nuclear reactions such as: (n,p) and (n,α) reactions. To provide an additional check on the systematic uncertainty two different cobalt sources have been used. The absolute activity for each gamma-ray together with the evaluated random and systematic uncertainty were reported by the laboratories. The actitivty correction for decay was performed using half life data either presented in the source certificate or from more recent literature.

Table 2

Source	E (KeV)	Efficiency	Random Error (%) 1 = 1	Systematic Error (%) 1 = 2
	1408.03	2.683E-04	0.4472	1.4908
	1112.00	3.284E-04	0.4872	1.4908
	1086.40	3.320E-04	0.7229	1.5060
Eu-152	964.00	3.681E-04	0.5161	1.4941
	867.40	4.031E-04	0.5954	1.4884
	778.90	4.315E-04	0.5562	1.5064
Bi-207	569.60	5.784E-04	0.3424	1.0027
	1332.50	3.383E-04	0.3843	1.0050
Co-60	1173.20	3.084E-04	0.3567	1.0052
(1)	1332.50	2.771E-04	0.3609	1.0104
Co-60	1173.20	3.089E-04	0.3561	0.8740
(2)	1332.50	2.783E-04	0.3593	0.8624

The calibration method employed here is the usual in routine activation analysis. Four standard sources, ~ 3 mm in diameter, were used to stablish the bare-point source gamma-ray efficiency curve for full energy peaks versus gamma-ray energy, at the position shown in Fig.1.A distance of ~ 20 cm between source and detector was chosen in order to avoid the problem of sum-coincidences. Activity measurements for the standard sources were achieved by counting the well known emmited gamma-rays with a Ge(Li) detector having an active volume of ~ 100 cm^3.Counting dead time corrections were small and they were deduced for each standard source count, using information recorded during these runs. In order to verify the reproducibility several independents runs were performed for each source. The final counting for each gamma-ray was obtained through weighted average of the equivalent quantities. The χ^2 normalized for all measurements were lower than 1, showing a consistency between the scatter

of the data and the assigned errors. Relative efficiencies determined as described above are presented in Table 2 and the evaluated error components appear there as well. The random uncertainty includes both statistic and reproducibility error of the calibration experiment plus, the random component in the quoted errors for intensity of a particular gamma-line in the standard sources. The systematic uncertainty includes only the systematic component quoted by the laboratories.

DATA ANALYSIS

As was commented previously, the goals of this data analysis are to obtain the best fitting to the calibration data using least squares method and the covariance matrix for the fitting parameter errors.We now turn our attention to determining first the covariance matrix for the set of data (efficiencies). As can be seen in Table 2, for the present case there are only two attributes which determine the overall uncertainties for the experimental data. The first one is the random error or statistic error (l=1) and the second one is the systematic error (l=2). One can assume the correlation level for the error components in the following form: no correlation for all random errors ($C_{ij1}=0$, $i \neq j$); complete correlation ($C_{ij2}=1$) for the systematic errors among gamma-rays emmited from the same source and no correlation ($C_{i'j'2}=0, i' \neq j'$) for the systematic errors among gamma-rays emmited from different sources. It is fair to assume in this example (and in most of the practical situations) that there is no correlation between the standard errors for sources from two distinct laboratories. The correlation matrix for each attribute can now be easily generate and the results are presented in Tabie 3.

Table 3

Correlation Matrices

1 = 1 (Random Error)	1 = 2 (Systematic Error)
1	1
0 1	1 1
0 0 1	1 1 1
0 0 0 1	1 1 1 1
0 0 0 0 1	1 1 1 1 1
0 0 0 0 0 1	1 1 1 1 1 1
0 0 0 0 0 0 1	0 0 0 0 0 0 1
0 0 0 0 0 0 0 1	0 0 0 0 0 0 1 1
0 0 0 0 0 0 0 0 1	0 0 0 0 0 0 0 0 1
0 0 0 0 0 0 0 0 0 1	0 0 0 0 0 0 0 0 1 1
0 0 0 0 0 0 0 0 0 0 1	0 0 0 0 0 0 0 0 0 1
0 0 0 0 0 0 0 0 0 0 0 1	0 0 0 0 0 0 0 0 0 1 1

The covariance matrix for the data point could be readly obtained (see Appendix A) making use of Eq. 1. However one could be interested to consider only one data set for the cobalt sources. Despite not to be relevant for the present analysis, this is a very usual procedure employed in experimental research which merits some discussion. It is common in any experiment that a particular parameter (for example cross section)be measured several times at the same energy, in order to provide an opportunity for identifying sources of systematic uncertainties. However, for reporting purposes or data analysis, it may be desirable to average all equivalent quantities. In this case, it becomes necessary to generate the corresponding covariance matrix for the resulting data set. In the present analysis it is desirable do convert 4 efficiency data, obtained with the cobalt sources, in only 2 data and to obtain the respective covariance matrix. According to SMITH (12) the procedure consists in collapsing the 4 data points on 2 by a proper weighted averaging technique based on the least-squares method. Furthermore, it was also pointed out that the elements of the covariance matrix for the collapsed set are linear combinations of elements from the covariance matrix of the origi-

nal data. These procedures when applied to the present analysis produce the results which are shown in Table 4. As can be seen the small χ^2 for each case shows once more that there is a consistency in the assumed errors.

Table 4

Weighted average and respective correlation matrix for the gamma-rays emmited from the cobalt sources.

Gamma-Line (KeV)	Efficiency	Error (%)	Average Value	Error (%)	Chi-Square	Correlation Matrix	
1) 1173.2	3.084E-04	1.066					
3) 1173.2	3.089E-04	0.962	1) 3.087E-04	0.71	0.013	1.0	
2) 1332.5	2.771E-04	1.052					
4) 1332.5	2.783E-04	0.934	2) 2.778E-04	0.70	0.094	0.868	1.0

Now one wishes to obtain the covariance matrix for the remaining 10 data points. The procedure is the same as discussed above and Table 5 shows the results produced by CALCOV. The covariance matrix presented in this table was tested for positive definiteness property using the program MATXTST. The matrix was classified as positive definite and so it is consisten with the physical reality. This covariance matrix has been used in the fitting to the experimental data which will be next discussed.

Table 5

The Covariance Matrix (%)

2.422									
2.222	2.460								
2.245	2.245	2.791							
2.227	2.227	2.250	2.499						
2.219	2.219	2.242	2.224	2.570					
2.246	2.246	2.269	2.251	2.242	2.579				
0.0	0.0	0.0	0.0	0.0	0.0	1.123			
0.0	0.0	0.0	0.0	0.0	0.0	1.008	1.158		
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.504	
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.432	0.490

The Parameter Total Errors (%)

1.556	1.568	1.671	1.581	1.603	1.606	1.060	1.076	0.710	0.700

The Correlation Matrix (Multiplied by 1000)

1000									
910	1000								
863	856	1000							
905	898	852	1000						
889	882	837	877	1000					
898	891	845	886	870	1000				
0	0	0	0	0	0	1000			
0	0	0	0	0	0	883	1000		
0	0	0	0	0	0	0	0	1000	
0	0	0	0	0	0	0	0	868	1000

It has been found (13) that the relationship between efficiency ε_f and gamma-ray energy E can be approximated by the expression:

12) $$\ln(\varepsilon_f) = \sum_{i=1}^{n} p_i (\ln E)^{i-1}$$

In order to determine the coefficients p_i, the expression above was fitted to calibration data using least squares method, discussed in the section Formalism. A computer program has been developed to perform this fitting, using the subroutine LLSF presented in ref. 13. The best curve obtained was for n=2 and the result is shown in Fig. 2. Higher order fits lead to a poorer (larger normalized χ^2) or to an unstability due to computer precision limitations. The equation representing the best fitting is given by:

$$13) \quad \ln(\varepsilon_f) = (-2.111 \pm 0.071) - (0.844 \pm 0.010) \ln(E)$$

and the covariance ($\bar{V}_p$) and correlation ($\bar{C}_p$) matrices for the fitting parameters are:

$$14) \quad \bar{V}_p = \begin{matrix} 5.035E\text{-}03 & \\ -7.104E\text{-}04 & 1.018E\text{-}04 \end{matrix} \qquad \bar{C}_p = \begin{matrix} 1 & \\ -0.992 & 1 \end{matrix}$$

The χ^2 normalized obtained for n=2 was near unity and thus the solution represented by Fig.2 or Eq. 13 may be considered as satisfactory.

The next step now is how to obtain the error ($\Delta\varepsilon_f$) in the relative efficiency determination for a gamma-ray with a particular energy, using the results of Eqs. 13 and 14. This particular case was also discussed in the section Formalism and the following expression represents the solution,

$$15) \quad (\Delta\varepsilon_f)^2 = (\partial Z/\partial p_1 \quad \partial Z/\partial p_2)\ (\bar{V}_p)\ \begin{matrix} (\partial Z/\partial p_1) \\ (\partial Z/\partial p_2) \end{matrix}$$

where $Z = \ln(\varepsilon_f)$ at a particular energy E. For instance, if one wishes to know the relative efficiency for the gamma-ray emmited from Na-24 (E = 1368.6 KeV), the following result is obtained:

16) ε_f (1368.6) = 0.27E-03 ± 0.25E-05

For the experimental example discussed in this paper a polynomial of degree two was obtained as the best fitting to the calibration data and thus, the solution of Eq. 15 is very simple. For higher order polynomials this solution is not so trivial and a small FORTRAN program named EFFIC has been developed to handle this problem. Details on the computational procedure of this program is presented in the Appendix B.

CONCLUSION

With the high development level of the Nuclear Physics and its technological application there is, at present moment, a greater need to become concerned with accuracy and proper estimation and representation of errors. In this paper we have tried to show the covariance matrix and least-squares methodology applied to a simple example of Nuclear Physics experiments. It is our concern that the statistical methods discussed here seem to provide a reasonably adequate approach for meeting this need. It is worthwhile for scientists working on Nuclear Physics or other areas to be knowledgeable regarding these methods.

REFERENCES

1. PEREY, F.G. "Estimated uncertainties in nuclear data - an approach". In: Shuack, R.R. & Bowman, C.D. "Nuclear cross sections and technology. Washington D.C., U.S. National Bureau of Standards, 1975, p. 842.

2. PEREY, F.G. "Covariance matrices of experimental data". In: Neutron physics and nuclear data: proceeding of the International Conference ... held in Harwell, U.K., 1978.

3. SMITH, D.L. "Covariance matrices and applications to the field of nuclear data", Idaho Falls, Argonne National Lab., 1981. (ANL/NDM-62).

4. MANNHART, W. "A small guide to generating covariances of experimental data". Braunschweig (Germany, F.R.), Physikalisch-Technische Bundesanstalt, 1981. (PTB-FMRB-84).

5. MANNHART, W. "Progress in integral data and their accuracy: average neutron cross sections in the californium-252 benchmark field". Nucl. Sci. Eng., 77:40, 1981.

6. CORCUERA, R.P. "Covariance matrices in evaluated nuclear data". Vienna, International Atomic Energy Agency, 1985. INDC(BZL)-16/GI INT(85)-2.

7. MUCCIOLO, E.R. & HELENE, O. "Energy measurement of gamma-rays from Ba-133 decay: consistency with cascade cross-over relations". Nucl. Instrum. Methods Phys. Res., A256:153-160, 1987.

8. SMITH, D.S. & GERALDO, L.P. "A vector model for error propagation". Am. J. Phys. (To be published).

9. COATES, M.S.; GAYTHER, D.B.; JAMES, G.D.; MOXON, M.C.; PATRICK, B.H.; SOWERBY, M.G.; SYME, D.B. "Can we do more to achieve accurate nuclear data". In: BOECKHOFF, K.H. & REIDEL, D. Nuclear data for science and technology. Dordrecht, Holland, Publishing Company, 1983. p. 977.

10. GERALDO, L.P. & SMITH, D.L. "Some thoughts on positive definiteness in the consideration of nuclear data covariance matrices" Idaho Falls, Argonne National Lab., 1988. (ANL/NDM-104).

11. PEELE, R.W. "Requirements on experiment reporting to meet evaluation needs". In: POENITZ, W.P. & SMITH, A.B. Fast neutron fission cross sections of U-233, U-238, Pu-239: proceedings of the NEANDC/NEACRP specialists meeting on ... held on Idaho Falls, Argonne National Lab., June 28-30, 1976. p. 421.

12. SMITH, D.L. "Generation of the covariance matrix for a set of nuclear data by collapsing a larger parent set through the weighted averaging of equivalent data points. Nucl. Instrum. Methods in Phys. Res., A257:365, 1987.

13. SMITH, D.L. "Non-evaluation applications for covariance matrices". Idaho Falls, Argonne National Lab., 1982. (ANL/NDM-67).

APPENDIX A

CALCOV PROGRAM (FOR PC OR COMPATIBLE)

A FORTRAN program named CALCOV has been written to construct covariance matrix for a data set x_i (i=1,n), according to the following relationship:

$$V_{ij} = \sum_{l=1}^{L} C_{ijl} \; E_{il} \; E_{jl} \quad \text{where}$$

C_{ijl} - are the microcorrelations

E_{il}, E_{jl} - is the error introduced by each attribute l.

Before running CALCOV it is necessary to create the input and output files. In the input fii the following sequence of data and format must appear:

1st line: N, LL format 2I5 where,

N = n, is the number of data points and

LL= L, is the number of attributes.

next lines: I = 1,N

(E(I,L), L =1,LL) format 7E10.4 where,

E(I,L) = E_{il}, are the partial errors.

The output file is used by CALCOV to store the results of the calculus, i.e., the covariance matrix, the correlation matrix and the total errors for the data set.

Running CALCOV

All the prompts announced by the program are self explanatory as is evident from the example provided in this appendix. The first procedure of CALCOV is to help the user to input all the correlation matrices for the attributes. For this endeavour it is announced the attribute number being processed and, the user is prompted for the majority corre-

lation on which all the matrix elements will be setted. Next the user can perform any change on the matrix according to the following options:

S = SINGLE, T = TRIANGLE, B = BLOCK, E = EXIT, P = PRINT where,

S = enter single element (C_{iji}) - to change or over-reading only one element of the matrix

T = enter triangular block - to change or over-reading any triangular block of elements in the matrix.

B = enter rectangular block - to change or over-reading any rectangular block of elements in the matrix.

P = prints the current correlation matrix.

E = exits and proceeds to the next attribute I.

After the construction of all correlation matrices, the covariance matrix is generated together with the final correlation matrix (with elements multiplied by 1000) and the total errors. All these informations are stored in the output file created by the user. After the listing of the code it is presented the results of running CALCOV, for the example discussed in this paper.

```
      PROGRAM CALLOV
      COMMON X(45,45),          (illegible)
      INTEGER IC(45,45)
      DATA IS,IT,IA,IE,IF,IN/'S','T','A','E','F','N'/
 9999 WRITE(*,9998)
 9998 FORMAT(1X,'CALLOV PROGRAM ')
      PAUSE
C
C     INPUT DATA
C
      WRITE(*,9997)
 9997 FORMAT(1X,'DATA INPUT: ')
      OPEN(3,FILE=' ',STATUS='OLD ')
      READ(3,2) N,LL
    2 FORMAT(I6,I5)
      DO 3 I=1,N
    3 READ(3,4) (E(I,L),L=1,LL)
    4 FORMAT(7E10.4)
C
C     SET COVARIANCE MATRIX TO ZERO
C
      DO 10 I=1,N
      DO 10 J=1,I
   10 C(I,J)=0.0
C
C     START ATTRIBUTE LOOP
C
 1000 WRITE(*,11)
   11 FORMAT(1X,'Do you intend to save the correlation matrix for any at
     1tribute Y/N ')
      READ(*,12,ERR=1001)IOP
   12 FORMAT(A1)
      IF(IOP .EQ. IN) GO TO 13
      WRITE(*,111)
  111 FORMAT(1X,'Please type the file name where the matrix will be stor
     1ed ')
      OPEN(5,FILE=' ',STATUS='NEW')
   13 WRITE(*,14)
   14 FORMAT(1X,'Do you want to input, from a file, the correlation matr
     1ix for any attribute? Y/N ')
      READ(*,12,ERR=17)IOP
      IF(IOP .EQ. IN) GO TO 15
      OPEN(4,FILE=' ',STATUS='OLD ')
   15 DO 1000 I=1,LL
C
C     SELECT ATTRIBUTE MAJORITY CORRELATION
C
      WRITE(*,20) L
   20 FORMAT(1X,'The Attribute Number is: ',I5)
  151 WRITE(*,201)
  201 FORMAT(1X,'Do you want to input the correlation matrix for this at
     1tribute (verify your input file)? Y/N ')
      READ(*,12,ERR=151)IOP
      IF(IOP .EQ. IN) GO TO 202
      DO 203 I=1,N
  203 READ(4,4,ERR=151) (D(I,J), J=1,I)
      GO TO 400
  202 WRITE(*,204)
  204 FORMAT(1X,'The Majority Correlation for attribute is:',I5)
  205 READ(*,4,ERR=203) DUM
      DO 21 I=1,N
      DO 21 J=1,I
```

```
   21 Q(I,J)=QLM
      GO TO 800
C
C     ALTER SELECTED CORRELATIONS D  INTERACTIVELY. SELECT 1=0 TO EXIT
C
  400 WRITE(*,401)
  401 FORMAT(1X,'OPTION S=SINGLE,I=INCREMENT,D=DONE,E=EXIT,P=PRINT')
  402 READ(*,403) IOP
  403 FORMAT(A1)
      IF(IOP.EQ.'S') GO TO 500
      IF(IOP.EQ.'I') GO TO 600
      IF(IOP.EQ.'D') GO TO 700
      IF(IOP.EQ.'P') GO TO 750
      IF(IOP.EQ.'E') GO TO 50
      GO TO 402
  500 WRITE(*,501)
  501 FORMAT(1X,'I,J (2I5) ')
      READ(*,2,ERR=500) I,J
      IF(I.LE.0) GO TO 500
      IF(I.GT.N) GO TO 500
      IF(J.GT.I) GO TO 500
  5011 WRITE(*,502)
  502 FORMAT(1X,'Q (E10.4) ')
      READ(*,4,ERR=5011) QVAL
      Q(I,J)=QVAL
      GO TO 800
  600 WRITE(*,601)
  601 FORMAT(1X,'IMIN,IMAX (2I5)')
      READ(*,2,ERR=600) IMIN,IMAX
      IF(IMIN.LE.0) GO TO 600
      IF(IMIN.GT.N) GO TO 600
      IF(IMAX.LE.0) GO TO 600
      IF(IMAX.LT.IMIN) GO TO 600
      IF(IMAX.GT.N) GO TO 600
 6011 WRITE(*,502)
      READ(*,4,ERR=6011) QVAL
      DO 602 I=IMIN,IMAX
      DO 602 J=IMIN,I
  602 Q(I,J)=QVAL
      GO TO 800
  700 WRITE(*,601)
      READ(*,2,ERR=700) JMIN,JMAX
      IF(JMIN.LE.0) GO TO 700
      IF(JMIN.GT.N) GO TO 700
      IF(JMAX.LE.0) GO TO 700
      IF(JMAX.LT.JMIN) GO TO 700
      IF(JMAX.GT.N) GO TO 700
  701 WRITE(*,702)
  702 FORMAT(1X,'JMIN,JMAX (2I5)')
      READ(*,2,ERR=701) JMIN,JMAX
      IF(JMIN.LE.0) GO TO 701
      IF(JMIN.GT.N) GO TO 701
      IF(JMAX.LE.0) GO TO 701
      IF(JMAX.LT.JMIN) GO TO 701
      IF(JMAX.GT.N) GO TO 701
 7021 WRITE(*,502)
      READ(*,4,ERR=7021) QVAL
      DO 703 I=IMIN,IMAX
      DO 703 J=JMIN,JMAX
  703 Q(I,J)=QVAL
      GO TO 800
  750 WRITE(*,755) L
  755 FORMAT(1X,'Q(I,J) FOR L = ',I5)
```

```
      DO 760 I=1,N
  760 WRITE(*,101) (Q(I,J),J=1,I)
      GO TO 400
C
C     OVERRIDE TO SET ATTRIBUTE SELF CORRELATION TO UNITY
C
  900 CONTINUE
      DO 30 I=1,N
   30 Q(I,I)=1.0
      GO TO 400
C
C     UPDATE COVARIANCE MATRIX
C
   50 CONTINUE
      WRITE(*,51)
   51 FORMAT(1X,'Do you want to save the correlation matrix for this at
     tribute? Y/N ')
      READ(*,12,ERR=50)IOF
      IF(IOF .EQ. 1N) GO TO 53
      DO 52 I=1,N
   52 WRITE(6,4) (Q(I,J), J=1,I)
   53 DO 54 I=1,N
      DO 54 J=1,I
   54 V(I,J)=V(I,J)*Q(I,J)*E(I,L)*E(J,L)
C
C     END ATTRIBUTE LOOP
C
 1000 CONTINUE
C
C     OUTPUT COVARIANCE MATRIX, ERRORS AND CORRELATION MATRIX ON
C     OUTPUT FILE
C
      WRITE(*,99)
   99 FORMAT(1X,'PLEASE TYPE THE OUTPUT FILE NAME ')
      OPEN(5,FILE=' ',STATUS='OLD ')
      WRITE(5,98)
   98 FORMAT(1X,'Number of Data Points:  ')
      WRITE(5,2) N
      WRITE(5,97)
   97 FORMAT(1X,'The Covariance Matrix is: ')
      DO 100 I=1,N
  100 WRITE(5,101) (V(I,J),J=1,I)
  101 FORMAT(7E11.4)
      WRITE(5,96)
   96 FORMAT(1X,'The Parameter Errors are: ')
      DO 110 I=1,N
  110 ER(I)=SQRT(V(I,I))
      WRITE(5,101) (ER(I),I=1,N)
      WRITE(5,95)
   95 FORMAT(1X,'The Correlation Matrix is: ')
      DO 120 I=1,N
      DO 120 J=1,I
      IC(I,J)=1000*V(I,J)/ER(I)/ER(J)
      IF(I .EQ. J) IC(I,J)=1000
  120 CONTINUE
      DO 130 I=1,N
  130 WRITE(5,135) (IC(I,J),J=1,I)
  135 FORMAT(20I4)
C
      STOP
      END
```

RUNNING CALCOV

```
CALCOV PROGRAM
Pause.
Please press (return) to continue.
DATA INPUT: File name missing or blank - Please enter name

UNIT 3? CALCOV.INP
 Do you intend to save the correlation matrix for any attribute?Y/N
N
Do you want to input, from a file, the correlation matrix for any attribute ? Y
N
The Attribute Number is:      1

Do you want to input the correlation matrix for this attribute (verify your inpu
t file)? Y/N
N
The Majority Correlation for Attribute(F10.4)
0.0
OPTION S=SINGLE,T TRIANGLE,B=BLOCK,E=EXIT,P PRINT
E
Do you want to save the correlation matrix for this attribute? Y/N
N
The Attribute Number is:      2

Do you want to input the correlation matrix for this attribute (verify your inpu
t file)? Y/N
N
The Majority Correlation for Attribute(F10.4)
0.0
OPTION S=SINGLE,T=TRIANGLE,B=BLOCK,E=EXIT,P=PRINT
T
IMIN,IMAX (215)
    1    6
U (F10.4)
1.0
OPTION S=SINGLE,T=TRIANGLE,B=BLOCK,E=EXIT,P=PRINT
S
I,J (215)
    8    7
U (F10.4)
1.0
OPTION S=SINGLE,T=TRIANGLE,B=BLOCK,E=EXIT,P=PRINT
S
I,J (215)
   10    9
U (F10.4)
0.8486
OPTION S=SINGLE,T=TRIANGLE,B=BLOCK,E=EXIT,P=PRINT
E
Do you want to save the correlation matrix for this attribute? Y/N
N
PLEASE TYPE THE OUTPUT FILE NAME
UNIT 5? CALCOV.OUT
Stop - Program terminated.
```

```
    CALCOV.OUT
Number of Data Points:
  10
The Covariance Matrix is:
 .2422E+01
 .2222E+01  .2460E+01
 .2245E+01  .2245E+01  .2791E+01
 .2227E+01  .2227E+01  .2250E+01  .2499E+01
 .2219E+01  .2219E+01  .2242E+01  .2224E+01  .2570E+01
 .2246E+01  .2246E+01  .2269E+01  .2251E+01  .2242E+01  .2579E+01
 .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00  .1123E+01
 .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00  .1008E+01
 .1158E+01
 .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00
 .0000E+00  .5041E+00
 .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00  .0000E+00
 .0000E+00  .4317E+00  .4900E+00
The Parameter Errors are:
 .1556E+01  .1568E+01  .1671E+01  .1581E+01  .1603E+01  .1606E+01  .1060E+01
 .1076E+01  .7100E+00  .7000E+00
The Correlation Matrix is:
1000
 9101000
 863 8561000
 905 898 8521000
 889 882 837 8771000
 898 891 845 886 8701000
   0   0   0   0   0   01000
   0   0   0   0   0   0 88310000
   0   0   0   0   0   0   0   01000
   0   0   0   0   0   0   0   0 8681000
```

APPENDIX B

EFFIC PROGRAM (FOR PC OR COMPATIBLE)

In the relative efficiency calibration for a Ge(Li) detector the following expression has been used to fit the experimental data:

$$\ln (\varepsilon_f) = \sum_{i=1}^{n} p_i (\ln E)^{i-1} \quad \text{where,}$$

ε_f - are the relative efficiency data

E - are the gamma-ray energies

The coefficients p_i together with the respective covariance matrix $\bar{V}_p$ are determined through least-squares fitting, discussed with some detail in this paper. In this way, the relative efficiency for a gamma-ray with a particular energy can be readly obtained with the results of the fitting. A simple FORTRAN program named EFFIC has been developed to determine the error $(\Delta\varepsilon_f)$ for this relative efficiency, according to the expression:

$$(\Delta\varepsilon_f)^2 = \bar{T}^t \bar{V}_p \bar{T} \quad \text{where,}$$

the symbol "t" denotes matrix transposition and $\bar{T}$, the transformation matrix, is given by:

$$\bar{T} = \begin{pmatrix} \partial Z/\partial p_1 \\ \partial Z/\partial p_2 \\ \partial Z/\partial p_i \end{pmatrix} \quad \text{where } \bar{Z} = \ln (\varepsilon_f) \text{ at the particular energy E.}$$

The program besides the error EFF = $\Delta\varepsilon_f$ also calculates the relative efficiency EF = ε_f at the required energy. All the prompts of EFFIC are self explanatory as can be seen in the sample problem provided after the listing of the code. This sample problem represents the example, for the gamma-ray emmited from Na-24, discussed at the end of the section Data Analysis.

```
      DIMENSION V(10,10),S(10),T(10),P(10)
      REAL  E,SOMA,LNE,EZ,PROD,Z,EF,EEF
      WRITE(*,5)
    5 FORMAT(1X,'INPUT THE NUMBER OF PARAMETERS, I2')
      READ(*,10) N
   10 FORMAT(I2)
      WRITE(*,12)
   12 FORMAT(1X,'INPUT THE PARAMETERS VALUES,E11.4 ')
      DO 15 I=1,N
   15 READ(*,40) P(I)
      WRITE(*,18)
   18 FORMAT(1X,'INPUT THE ENERGY(KEV) VALUE, F7.2')
      READ(*,20) E
   20 FORMAT(F7.2)
      WRITE(*,25)
   25 FORMAT(1X,'INPUT THE PARAMETER COV.MATRIX, 6E11.4')
      DO 30 I=1,N
   30 READ(*,40)  (V(I,J),J=1,I)
   40 FORMAT(6E11.4)
      DO 50 I=1,N
      DO 50 J=1,N
   50 V(I,J)=V(J,I)
      LNE=LOG(E)
      DO 60 I=1,N
   60 T(I)=LNE**(I-1)
      DO 70 J=1,N
      SOMA=0.0
      DO 70 I=1,N
      SOMA=SOMA + V(I,J)*T(I)
   70 S(J)=SOMA
      DO 75 J=1,N
   75 WRITE(*,40) S(J)
      PROD=0.0
      DO 80 I=1,N
   80 PROD=PROD + S(I)*T(I)
      EZ=SQRT(PROD)
      WRITE(*,85) EZ
   85 FORMAT(1X,' EZ= ',E11.4)
      Z=0.0
      DO 90 I=1,N
   90 Z=Z + P(I)*T(I)
      EF=EXP(Z)
      EEF=EF*EZ
      WRITE(*,100) EF,EEF
  100 FORMAT(1X,' EF= ',E11.4,'    EFF= ',E11.4)
      STOP
      END
```

PUNNING EFFIC

```
INPUT THE NUMBER OF PARAMETERS, I2
 2
INPUT THE PARAMETER VALUES,E11.4
 2.[illegible]
 0.[illegible]
INPUT THE [illegible] VALUE, F7.2
[illegible]
INPUT THE PARAMETER COV.MATRIX, 4E11.4
[illegible]
 [illegible]104 0.0000104185
   [illegible]E 04
 [illegible]E 04
 [illegible]    [illegible]E 02
 [illegible]    .[illegible]E 03    EFF[illegible]    .[illegible]E 05
[illegible]    Program terminated.
```

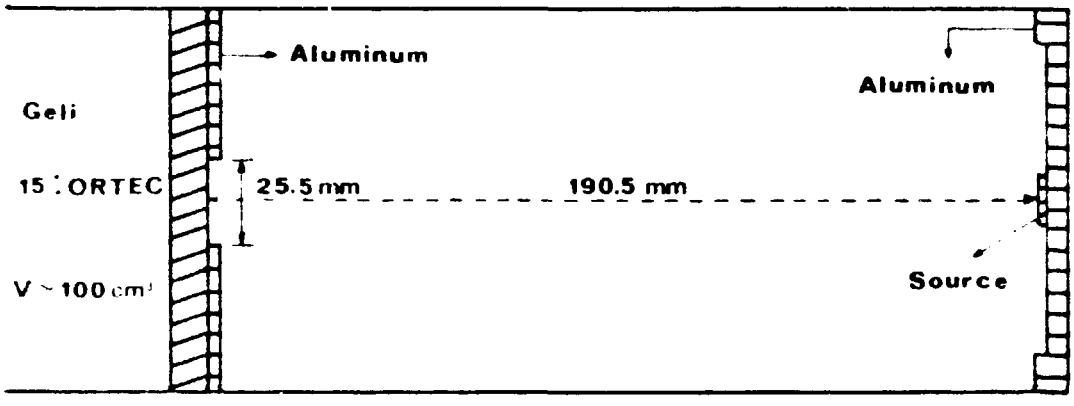

Aluminum
Aluminum
Geli
15 : ORTEC
25.5 mm
190.5 mm
V ~ 100 cm³
Source

FIG. 1

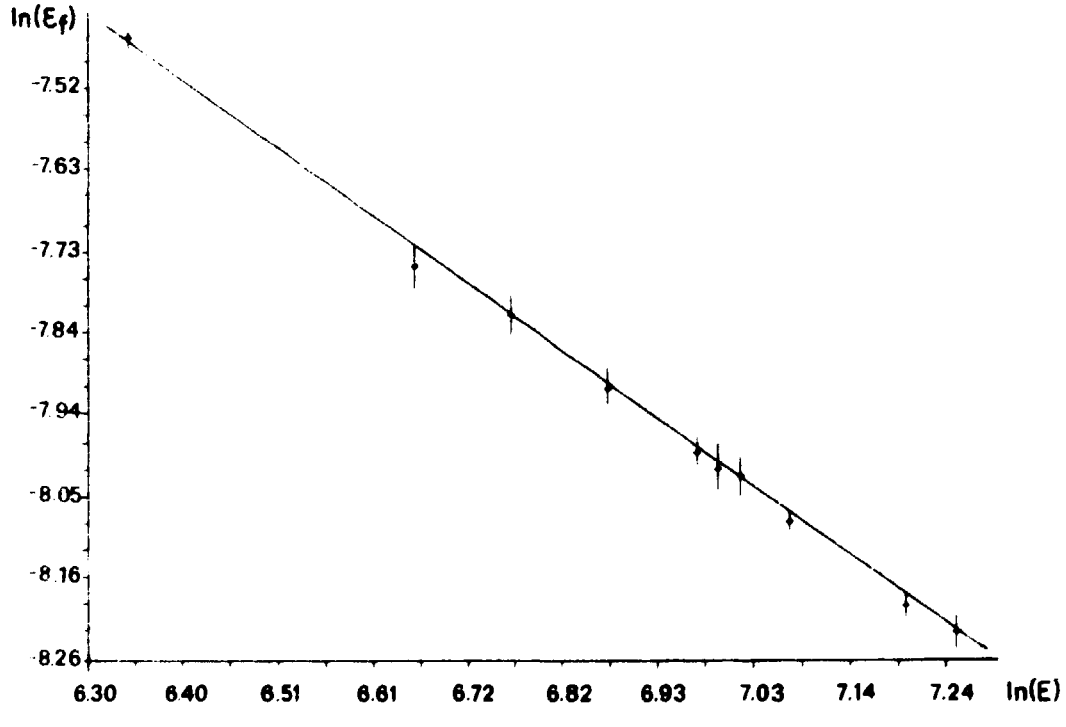

ln(Eγ)
-7.52
-7.63
-7.73
-7.84
-7.94
-8.05
-8.16
-8.26
6.30
6.40
6.51
6.61
6.72
6.82
6.93
7.03
7.14
7.24
ln(E)

FIG. 2