

Determination of Decay Constant by the Maximum Likelihood Method

Takeshi MUKOYAMA*

Received December 13, 1975

A method for analysis of decay data is presented. The method is based on maximum likelihood theory and so general that it is applicable to any other experimental distributions. The validity of the proposed method has been successfully tested against the artificial decay data. It is pointed out that the present method is more simple than the previous maximum likelihood methods.

I. INTRODUCTION

Since the discovery of the radioactivity, the measurement of decay constant (or nuclear life time) has been an important problem in nuclear physics. Various methods to analyze the experimental decay curves have so far been reported. The most commonly used method is to plot the logarithm of the background-subtracted counting rate versus time. The values thus obtained are fitted by least squares to a straight line, and the decay constant can be determined as the slope of this line. This method is very simple but the precision is poor, because the logarithmic transformation of the experimental data introduces large uncertainty.

Recently the influence of the environment on the disintegration rate of the radioactive nuclides has been studied by directly comparing the decay curves for the different chemical or physical conditions.¹⁻⁴⁾ In these experiments, the accuracy of the results depends often on the method used to extract the decay constant, because the change in the decay constants is very small. For this purpose, an exponential least-squares fit technique⁵⁾ has frequently used.

The standard least-squares method is based on finding the parameters which minimize chi-square:

$$\chi^2 = \sum_{i=1}^N W_i [Y_i - y(x_i; a_1 \cdots a_M)]^2, \quad (1)$$

where $y(x; a_1 \cdots a_M)$ is an M -parameter function with one independent variable x , Y_i is the measured quantities at the i -th data point x_i , and W_i is the weight associated with Y_i . In the case of decay curve, the function $y(x; a_1 \cdots a_M)$, exponential function plus constant, is not linear; hence, the solution should be obtained through iteration.

There exists a *maximum likelihood method* which is a different approach from Eq. (1). The maximum likelihood method consists of determining the maximum of the likelihood function defined in the normal manner.⁶⁾ This method has suc-

* 向山 毅: Laboratory of Nuclear Radiation, Institute for Chemical Research, Kyoto University, Kyoto.

cessfully been applied for the case of two-component decay curve with constant background.⁷⁾ For fitting of Gaussian to γ -ray peaks, Ciampi *et al.*⁸⁾ showed that this method furnishes less standard deviations than the least-squares method.

It is the purpose of the present paper to provide a new technique to obtain best fit of exponential function plus constant to the experimental decay data by the maximum likelihood method. Although the discussion is limited to the case of single-component decay curve, most of the arguments can also be applied to multi-component case, and we can easily extend the present method to any number of components. Moreover, this method is sufficiently general and readily applicable to most experimental distributions other than the exponential function.

II. THE MAXIMUM LIKELIHOOD METHOD

Let Y_i represent the number of counts accumulated in a given time interval at x_i . Because experimental counts have a Poisson distribution with a mean value m_i , the probability P_i of observing Y_i counts at x_i is

$$P_i = \exp(-m_i) (m_i)^{Y_i} / Y_i! . \quad (2)$$

The mean m_i is expressed as a function of undetermined parameters a_k ($k=1 \sim M$). If the probabilities, P_i , are uncorrelated, the joint probability is

$$L = \prod_{i=1}^N P_i , \quad (3)$$

where N is the total number of the data points. A maximum in L is obtained only when the parameters are the best values consistent with the function to be fitted.

Instead of maximization of L , it is more convenient to maximize the logarithm of L ;

$$\begin{aligned} W &= \ln L \\ &= \sum_{i=1}^N (Y_i \ln m_i - m_i) + \text{const.} \end{aligned} \quad (4)$$

In order to maximize W , Grard⁹⁾ proposed the method of steps (grid-search method). His method has the advantage that it does not require the first and the second derivatives of W . However, this method is time consuming.

In general, the maximum value of W can be obtained when all the first-order partial derivatives of W with respect to the parameters are equal to zero. For M parameters, the set of equations

$$\frac{\partial W}{\partial a_k} = 0; \quad k = 1 \sim M , \quad (5)$$

have to be solved simultaneously. In the case of decay data, where the function designed to fit is not linear with respect to a_k , it is not feasible to solve for the best values of a_k without iteration.

III. COMPUTATIONAL METHOD

From Eq. (4), $\partial W/\partial a_k$ can be written by

$$\frac{\partial W}{\partial a_k} = \sum_{i=1}^N (Y_i - m_i) / m_i \frac{\partial m_i}{\partial a_k}. \quad (6)$$

Denoting the value of the k -th parameter on the l -th iteration as a_k^l and expanding m_i in the first-order Taylor series, we obtain

$$m_i \simeq m_{i0} + \sum_{k=1}^M \left(\frac{\partial m_i}{\partial a_k} \right) (a_j^l; j=1, M) \Delta a_k, \quad (7)$$

where m_{i0} is the value of m_i calculated using a_k^l and

$$\Delta a_k = a_k^{l+1} - a_k^l.$$

When the parameter increments Δa_k are small, Eq. (6) can be approximated as

$$\frac{\partial W}{\partial a_k} \simeq \sum_{i=1}^N (Y_i - m_i) / m_{i0} \left(\frac{\partial m_i}{\partial a_k} \right) (a_j^l; j=1, M). \quad (8)$$

Substituting Eq. (7) into Eq. (8) and setting $\partial W/\partial a_k$ to zero [Eq. (5)], we obtain a set of M simultaneous linear equations with respect to Δa_k . These equations are directly soluble to give values of the parameter increments which may be used to improve the original estimates a_k^l . For the given initial estimates, the parameters a_k^l are improved with each iteration and the procedure is repeated until the convergence criteria are met.

Similar method has been developed by Orth *et al.*⁷⁾ They expanded $\partial W/\partial a_k$ of Eq. (5) in a first-order Taylor series by the use of the second-order partial derivatives of W . In this way, Eq. (5) reduces to M -simultaneous linear equations with respect to Δa_k . However, the number of the partial derivatives calculated in one iteration is $M(M+1)/2$ in their method, because they use the second-order partial derivative. On the other hand, the present method uses only the first-order derivative and the number of the derivatives is M . Therefore, the present method is more rapid and more simple than the method of Orth *et al.*

The variances associated with the fitted values a_k are, in general, not well defined. For large N , however, the likelihood function L approaches a Gaussian distribution at the neighbourhood of the maximum. In this case, the variances are well approximated as the square roots of the diagonal elements of the error matrix H^{-1} .⁶⁾ The matrix H is defined as

$$H_{jk} = (-\partial^2 W / \partial a_j \partial a_k) (a_j = a_j^l, a_k = a_k^l; j, k=1, M). \quad (9)$$

In the case of single-component decay, the decay curve is described by an exponential function plus a constant:

$$y = ae^{-\lambda x} + b, \quad (10)$$

where a , b , and λ are the parameters to be determined. Using this expression, the mean m_i is written by

$$m_i = ae^{-\lambda x_i} + b. \quad (11)$$

Three simultaneous equations are expressed as

$$\left. \begin{aligned} \sum_i (Y_i - m_i) \alpha_i / m_{i0} &= 0, \\ \sum_i (Y_i - m_i) \beta_i / m_{i0} &= 0, \\ \sum_i (Y_i - m_i) / m_{i0} &= 0, \end{aligned} \right\} \quad (12)$$

where α_i and β_i are the first-order partial derivatives of m_i with respect to a and λ , respectively. From Eq. (11), α_i and β_i are

$$\left. \begin{aligned} \alpha_i &\equiv e^{-\lambda x_i}, \\ \beta_i &\equiv -ax_i e^{-\lambda x_i}. \end{aligned} \right\} \quad (13)$$

Estimation of the uncertainties of the parameters is made using the error matrix, which is evaluated from the second-order partial derivatives:

$$\left. \begin{aligned} \frac{\partial^2 W}{\partial a^2} &= - \sum_i \alpha_i^2 Y_i / m_{i0}^2, \\ \frac{\partial^2 W}{\partial \lambda \partial a} &= - \sum_i \alpha_i [\beta_i Y_i / m_{i0}^2 + (Y_i / m_{i0} - 1) x_i], \\ \frac{\partial^2 W}{\partial b \partial a} &= - \sum_i \alpha_i Y_i / m_{i0}^2, \\ \frac{\partial^2 W}{\partial \lambda^2} &= - \sum_i \beta_i [\beta_i Y_i / m_{i0}^2 + (Y_i / m_{i0} - 1) x_i], \\ \frac{\partial^2 W}{\partial b \partial \lambda} &= - \sum_i \beta_i Y_i / m_{i0}^2, \\ \frac{\partial^2 W}{\partial b^2} &= - \sum_i Y_i / m_{i0}^2. \end{aligned} \right\} \quad (14)$$

IV. TESTING THE METHOD

In order to examine the validity of the present method, artificial data was used. For testing purpose, it is better to use the artificial data than the experimental data, because we know exactly what the true parameters are.

The artificial data was stochastically generated by computer simulation. The counts Y_i at the time x_i are considered to be Poisson distributed. For large Y_i values, this distribution can be approximated by a Gaussian one. According to the Gaussian distribution law, the probability P_i of observing Y_i counts at the time x_i is written by

$$P_i = [1/(2\pi)^2\sigma_i] \exp [-(Y_i-m_i)^2/2\sigma_i^2], \quad (15)$$

where m_i is the mean of Y_i counts and σ_i^2 is the variance associated with Y_i . In the present case, m_i and σ_i^2 are given by Eq. (11).

The random numbers with the Gaussian distribution, S_i , can be generated by solving the equation:¹⁰⁾

$$\int_{-\infty}^{S_i} P_i dY_i = R_i. \quad (16)$$

Here, R_i is a random number which is uniformly distributed on the interval [0, 1]. Substituting Eq. (15) into Eq. (16), we obtain

$$R_i = \frac{1}{2} \{1 + \text{erf}[(Y_i-m_i)/2^{1/2}\sigma_i]\}, \quad (17)$$

where $\text{erf}[x]$ is the error function.

For given values of a , b , and λ , the values Y_i at x_i are obtained from the uniform random numbers R_i by solving Eq. (17). For this purpose, the table of the error function is stored in the computer and the solution is obtained by the inverse interpolation. In order to reproduce the discrete character of the decay data, the integer number nearer to Y_i is taken to be the counts at x_i .

The artificial decay data was generated using the FACOM 230-75 computer of the Data Processing Center of Kyoto University. The parameters were chosen as

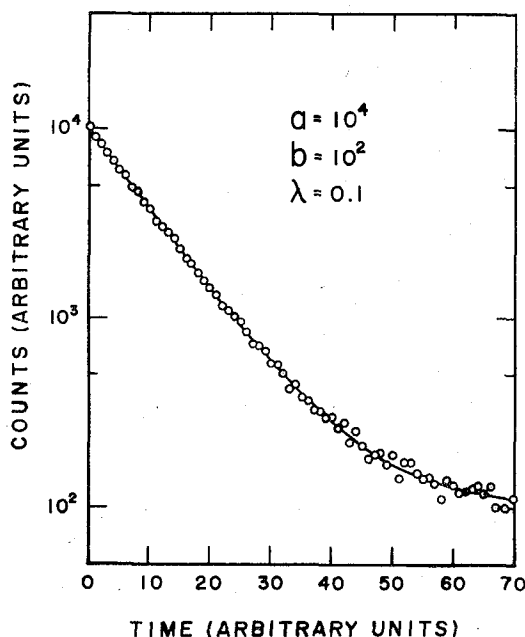


Fig. 1. Artificial decay data used to test the present method. The open circles are the pseudo-experimental values generated by the computer and the solid curve represents the fitted values by the present method.

$a=10000$, $b=100$, and $\lambda=0.1$. In Fig. 1, the pseudo-experimental decay data thus obtained is shown together with the values estimated by the maximum likelihood method. For this data, the estimates determined from the present method were: $a=9967\pm 46$, $b=102.2\pm 2.7$, and $\lambda=0.1002\pm 0.0004$. This result indicates that the present method can well reproduce the parameters used in the synthesis of the artificial decay curve.

It is noted that the good initial estimates are crucial in the present method. When the initial parameter estimates are poor, the iteration does not converge. This is because the approximation in Eq. (8) is invalid for such a case.

V. CONCLUSION

A new method to estimate the parameters for the decay data, based on the maximum likelihood theory, has been presented. The present method can yield the reasonable decay constant and is more simple than the method of Orth *et al.*

A disadvantage is its need of the good initial estimates. However, the situation is common to other iterative methods.

Extensions of the present method to multi-component case are derived with ease. Furthermore, its strategy is, in principle, applicable to most forms of experimental distributions other than the decay data; for example, Mössbauer spectra (Lorentzian) and γ -ray peaks (Gaussian).

ACKNOWLEDGMENT

The author wishes to express his thanks to Professor S. Shimizu for his encouragement throughout the present work.

REFERENCES

- (1) P. Küegsegger and W. Künding, *Phys. Letters*, **39B**, 620 (1973); *Helv. Phys. Acta*, **46**, 165 (1973).
- (2) M. Ulrickson, R. Hensler, D. Gordon, and N. Benczer-Koller, *Phys. Rev. C*, **9**, 326 (1974).
- (3) R. N. Verma and G.T. Emery, *Phys. Rev. B*, **9**, 3666 (1974).
- (4) P. Roggwiler and W. Künding, *Phys. Rev. B*, **11**, 4179 (1975).
- (5) For example, see "Applications of Computers to Nuclear and Radiochemistry", ed. by G.D. O'Kelly, NAS-NS 3107, U.S. AEC, 1962.
- (6) J. Orear, "Notes on Statistics for Physicists," UCRL-8417, 1958.
- (7) P. H. Orth, W.R. Falk, and G. Jones, *Nucl. Instr. and Meth.*, **65**, 301 (1968).
- (8) M. Ciampi, L. Daddi, and V. D'Angelo, *Nucl. Instr. and Meth.*, **66**, 102 (1968).
- (9) F. Grand. *Nucl. Instr. and Meth.*, **34**, 242 (1965).
- (10) "The Monte Carlo Method," ed. by Yu. A. Schreider, Pergamon Press, New York, 1962, p. 314.