Rapport BIPM-79/12

Simultaneous curve fitting

by Jörg W. Müller

Bureau International des Poids et Mesures, F-92310 Sèvres

Summary

The practical application of the well-known beta-gamma coincidence method for standardizing a radionuclide requires an extrapolation to a beta counting efficiency one. This is usually accomplished by fitting polynomials to the measured data. If there are several decay branches which can be separated by the use of appropriate gamma windows, the measurements, supposed to be independent, belong to several curves. The method of adjustment described in this report assures automatically that the polynomials, fitted simultaneously, have a common intersection for beta efficiency one, leading therefore to a unique value for the activity and its statistical uncertainty. The usefulness of the new fitting method is illustrated by an application to activity measurements of ^{134}Cs performed in the framework of a recent international comparison organized by BIPM.

1. Introduction

In any standardization of a radionuclide by means of the beta-gamma coincidence method, there is a step which has to be done empirically. This concerns a usually (but not always) minor correction which is due to the residual gamma sensitivity of the beta counter, to conversion electrons or to a complex decay-scheme branching. These rather ill-known efficiency-dependent effects are normally corrected for by a graphical extrapolation of the measurements to a beta efficiency ε_β of 100 %.

The traditional approach (for details see e.g. [1]) consists in plotting the quantity $y(x) = N_\beta \cdot N_\gamma / N_c$ as a function of, for instance, $x = (N_\gamma / N_c) - 1$, where N_β, N_γ and N_c are the empirical count rates corrected for background, dead time and accidental coincidences. Since $x = 0$ corresponds to $\varepsilon_\beta = N_c / N_\gamma = 1$, the activity N_o of the source measured is then given by the extrapolated value

$$N_o = y(0) \; . \qquad (1)$$

Let us now suppose that there are several decay branches which can be separated experimentally, for example by an appropriate setting of the gamma window. Each branch then gives a set of experimental values of y and x which lie on a different curve (see Fig. 1).

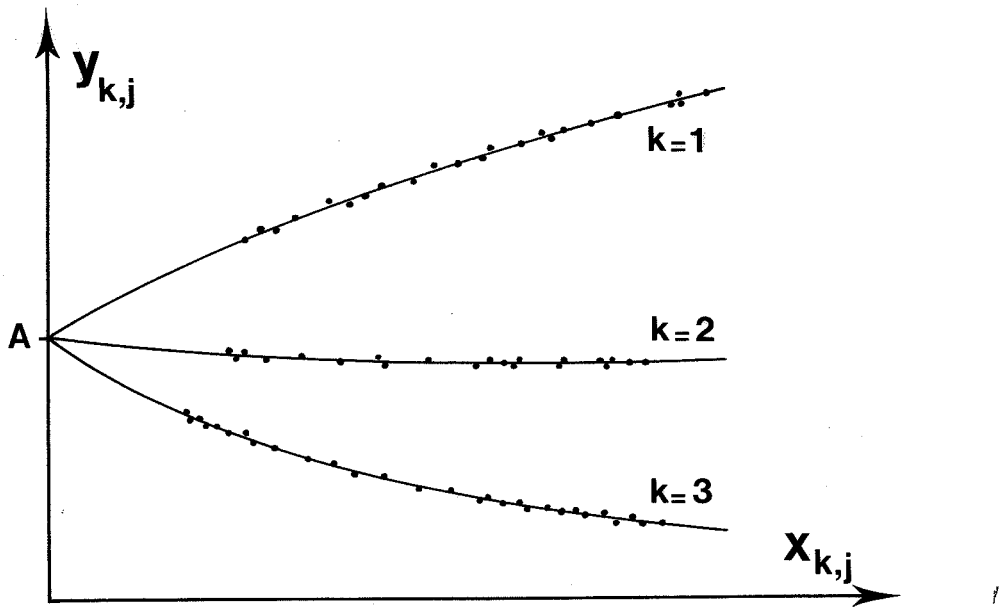

Figure 1 - Schematic plot of the apparent activities $y_{k,j}$ as a function of the corresponding "inverse" beta efficiencies $x_{k,j}$, for K = 3 gamma-channel gates. The common extrapolated value A is the activity looked for.

Since the activity N_o of the source should not depend on the decay branch chosen, the K curves are expected to have a common intersection at $x = 0$.

As the exact form of the different curves is unknown, it is usual to express them in the form of polynomials with coefficients to be adjusted. Thus, for K measured branches (k = 1, 2, ..., K) and polynomials for instance up to third order, the theoretical apparent activities y are expected to be given by the following set of equations

$$y_k(x_{k,j}) = A + B_k \cdot x_{k,j} + C_k \cdot x_{k,j}^2 + D_k \cdot x_{k,j}^3 \quad , \tag{2}$$

where j stands for a given measurement. Let us further assume that for branch k we have taken J_k empirical data of the form $Y_{k,j} \pm s_{k,j}$, where $s_{k,j}$ are the corresponding uncertainties (e.g. estimated standard deviations). Neglecting the uncertainty in $x_{k,j}$, we may attribute to each point in Fig. 1 a statistical weight which can be chosen as $g_{k,j} = s_{k,j}^{-2}$. For a simple approximate method to account for the correlation which in reality exists between the measured values of $Y_{k,j}$ and $x_{k,j}$ see [2].

2. Simultaneous least-squares fit

Application of the usual least-squares method leads to the condition that the sum of the weighted squared deviations, that is

$$Q = \sum_{k=1}^{K} \sum_{i=1}^{J_k} g_{k,i} \cdot v_{k,i}^2 \quad , \tag{3}$$

where $v_{k,i} = y_k(x_{k,i}) - Y_{k,i}$ are the residuals,

must be a minimum.

As is well known, this corresponds to the requirement that the partial derivatives of Q with respect to the parameters looked for vanish, hence

$$\frac{\partial Q}{\partial A} = 0, \qquad \frac{\partial Q}{\partial B_k} = 0, \qquad \frac{\partial Q}{\partial C_k} = 0 \qquad \text{and} \qquad \frac{\partial Q}{\partial D_k} = 0 \; . \tag{4}$$

This yields the 3K + 1 normal equations. Let us write them now more explicitly.

Thus, for instance $\frac{\partial Q}{\partial A} = 0$ corresponds to

$$\sum_{k=1}^{K} \sum_{i=1}^{J_k} g_{k,i} \cdot v_{k,i} = \sum_k \sum_i g_{k,i} \left(A + B_k x_{k,i} + C_k x_{k,i}^2 + D_k x_{k,i}^3 - Y_{k,i}\right) = 0 \; .$$

If we denote summation over the J_k measurements of a given branch by square brackets, for example

$$\sum_{i=1}^{J_k} g_{k,i} \cdot x_{k,i}^2 \equiv \left[g_k x_k^2 \right] \quad , \quad \text{etc.,}$$

we can also write

$$A \sum_k \left[g_k \right] + \sum_k B_k \left[g_k x_k \right] + \sum_k C_k \left[g_k x_k^2 \right] + \sum_k D_k \left[g_k x_k^3 \right] = \sum_k \left[g_k Y_k \right] .$$

Similarly, a condition like $\frac{\partial Q}{\partial C_k} = 0$ leads to

$$\sum_{i=1}^{J_k} g_{k,i} \cdot v_{k,i} \cdot x_{k,i}^2 = \sum_i g_{k,i} \left(A + B_k x_{k,i} + C_k x_{k,i}^2 + D_k x_{k,i}^3 - Y_{k,i}\right) x_{k,i}^2 = 0 \; ,$$

and application of the sum notation then gives

$$A \left[g_k x_k^2 \right] + B_k \left[g_k x_k^3 \right] + C_k \left[g_k x_k^4 \right] = \left[g_k x_k^2 Y_k \right] .$$

Therefore, the normal equations (4) for the present problem now take the form

$$
\begin{aligned}
A \sum_k \left[g_k\right] + \sum_k B_k \left[g_k x_k\right] + \sum_k C_k \left[g_k x_k^2\right] + \sum_k D_k \left[g_k x_k^3\right] &= \sum_k \left[g_k Y_k\right] \\
A \left[g_k x_k\right] + B_k \left[g_k x_k^2\right] + C_k \left[g_k x_k^3\right] + D_k \left[g_k x_k^4\right] &= \left[g_k x_k Y_k\right] \\
A \left[g_k x_k^2\right] + B_k \left[g_k x_k^3\right] + C_k \left[g_k x_k^4\right] + D_k \left[g_k x_k^5\right] &= \left[g_k x_k^2 Y_k\right] \\
A \left[g_k x_k^3\right] + B_k \left[g_k x_k^4\right] + C_k \left[g_k x_k^5\right] + D_k \left[g_k x_k^6\right] &= \left[g_k x_k^3 Y_k\right] ,
\end{aligned}
\qquad (5)
$$

where each of the last three lines corresponds to K equations.

Contrary to previous practice, the adjustment has now to be made simultaneously for all K branches since they are coupled to each other by the imposed condition to have a common point of intersection at $x = 0$.

Since for three branches the above system consists already of ten equations, it will be obvious that its solution requires the application of some powerful method, for instance in the form of the programs available as standard subroutines for a given type of computer. For our IBM 1130 machine, the program SIMQE is appropriate.

It is practical for this purpose to rewrite the set of simultaneous linear equations (5) in matrix form, for instance as

$$M \cdot X = Z , \qquad (6)$$

where

M = matrix of the abscisae (of size n by n),

X = vector of the wanted coefficients (length n),

Z = vector of the measurements (length n),

with $n = 3K + 1$ (for third order in x).

When written more explicitly, the square matrix M consists of the arrangement of elements as given in Table 1a (the subdivision by dashed lines is only for clarity).

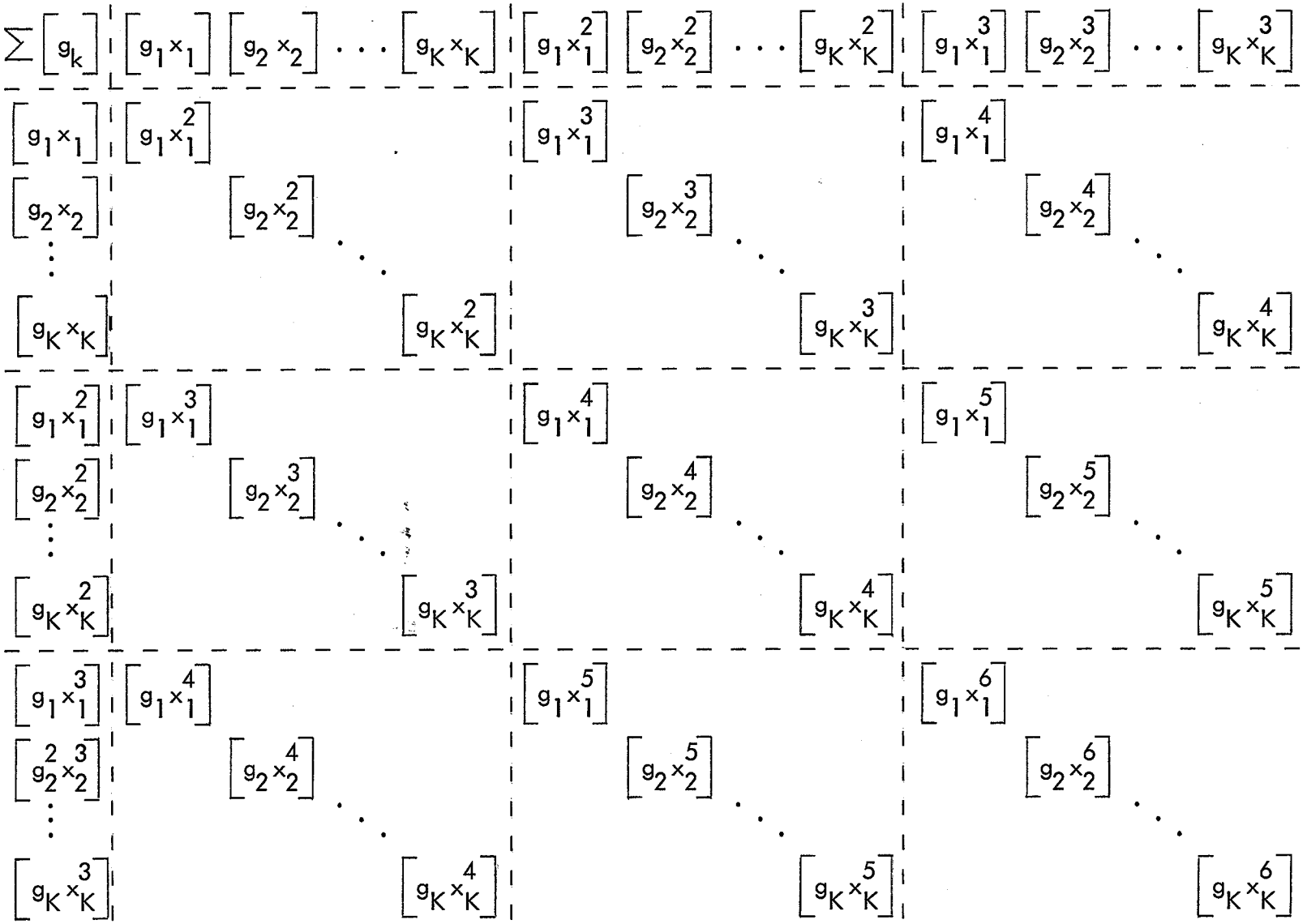

Table 1a - Arrangement of the elements of the square matrix M.
Note that all the elements outside the "diagonals" indicated vanish and that the matrix M is symmetric.

For the column vectors X and Z the elements are

$$X = \begin{pmatrix} A \\ \hline B_1 \\ \vdots \\ B_K \\ \hline C_1 \\ \vdots \\ C_K \\ \hline D_1 \\ \vdots \\ D_K \end{pmatrix}, \qquad Z = \begin{pmatrix} g_k Y_k \\ \hline g_1 Y_1 x_1 \\ \vdots \\ g_K Y_K x_K \\ \hline g_1 Y_1 x_1^2 \\ \vdots \\ g_K Y_K x_K^2 \\ \hline g_1 Y_1 x_1^3 \\ \vdots \\ g_K Y_K x_K^3 \end{pmatrix}.$$

as can be verified by a comparison with (5).

In view of the repeated occurrence of some sums as elements of M or of Z, it will be useful for their implementation into the computer program to form previously the following sums

$$\alpha = \sum_{k=1}^{K} \sum_{j=1}^{J_k} g_{k,j}$$

$$\beta = \sum_{k=1}^{K} \sum_{j=1}^{J_k} g_{k,j} \cdot Y_{k,j} \qquad (7a)$$

and, for an adjustment of up to order R in the variable x, with k = 1, 2, ..., K,

$$\gamma(k,s) = \sum_{j=1}^{J_k} g_{k,j} \cdot x_{k,j}^s \,, \qquad \text{with } s = 1, 2, \ldots, 2R \,,$$

$$\delta(k,r) = \sum_{j=1}^{J_k} g_{k,j} \cdot Y_{k,j} \cdot x_{k,j}^r \,, \qquad \text{with } r = 1, 2, \ldots, R \,. \qquad (7b)$$

This corresponds to $2 + 3\,R \cdot K$ summations (hence 38 for $K = 4$ and $R = 3$) which are automatically performed in the program. They use only the experimentally available data and give all the elements needed for M and Z.

The elements of the matrix M can now be written in the new variables as in Table 1b. Similarly Z is given by the elements β and $\delta(k,r)$.

The computer program is so arranged as to first solve the normal equations (with a fixed number $K \leqslant 4$ of branches) for a linear fit. Then M is inverted and the new n diagonal elements of M^{-1} are stored as they will be needed for the evaluation of the statistical uncertainties of the adjusted coefficients (see below). This procedure is then repeated for a second-order and finally for a third-order fit. It will be obvious that prior to each new adjustment the quantities M and Z must be dimensioned correspondingly and that the respective elements have to be put at their appropriate memory places. The program takes automatically care of all these details.

The evaluation of the estimated standard deviations for the adjusted parameters A, B_k, ... is done according to the classical prescription (see for instance [3] or [4] for the mathematical background). For a given fit, one calculates numerically the quantity

$$Q \equiv \chi^2 = \sum_{k=1}^{K} \sum_{j=1}^{J_k} g_{k,j} \cdot v_{k,j}^2 \quad , \tag{3}$$

where in

$$v_{k,j} = (A + B_k \cdot x_{k,j} + \ldots) - Y_{k,j}$$

the values of the parameters obtained in the fit are used. With

$$\sigma^2 = \frac{\chi^2}{\sum_{k=1}^{K} J_k - K \cdot r - 1} \quad , \tag{8}$$

where r is the order of the polynomial, the estimated variances can be expressed by means of the diagonal elements of the inverted matrix M^{-1} as ($k = 1, 2, \ldots, K$)

$$\begin{aligned} s^2(A) &= \sigma^2 \cdot M^{-1}(1,1) \quad , \\ s^2(B_k) &= \sigma^2 \cdot M^{-1}(1+k,\ 1+k), \\ s^2(C_k) &= \sigma^2 \cdot M^{-1}(1+K+k,\ 1+K+k), \\ s^2(D_k) &= \sigma^2 \cdot M^{-1}(1+2K+k,\ 1+2K+k) \quad . \end{aligned} \tag{9}$$

For our computer the inversion of the matrix M is performed by the subroutine MINVE.

α	$\gamma(1,1)$	$\gamma(2,1)$	...	$\gamma(K,1)$	$\gamma(1,2)$	$\gamma(2,2)$	...	$\gamma(K,2)$	$\gamma(1,3)$	$\gamma(2,3)$	...	$\gamma(K,3)$
$\gamma(1,1)$	$\gamma(1,2)$				$\gamma(1,3)$				$\gamma(1,4)$			
$\gamma(2,1)$		$\gamma(2,2)$				$\gamma(2,3)$				$\gamma(2,4)$		
...			. . .				. . .				. . .	
$\gamma(K,1)$				$\gamma(K,2)$				$\gamma(K,3)$				$\gamma(K,4)$
$\gamma(1,2)$	$\gamma(1,3)$				$\gamma(1,4)$				$\gamma(1,5)$			
$\gamma(2,2)$		$\gamma(2,3)$				$\gamma(2,4)$				$\gamma(2,5)$		
...			. . .				. . .				. . .	
$\gamma(K,2)$				$\gamma(K,3)$				$\gamma(K,4)$				$\gamma(K,5)$
$\gamma(1,3)$	$\gamma(1,4)$				$\gamma(1,5)$				$\gamma(1,6)$			
$\gamma(2,3)$		$\gamma(2,4)$				$\gamma(2,5)$				$\gamma(2,6)$		
...			. . .				. . .				. . .	
$\gamma(K,3)$				$\gamma(K,4)$				$\gamma(K,5)$				$\gamma(K,6)$

Table 1b - Elements of the matrix M, written by means of the sums defined in (7).

It should be remembered that the adjusted coefficients A, B_k, ... are correlated among themselves. The corresponding covariances are given by the out-of-diagonal elements of the matrix M^{-1} and they would have to be taken into account for assigning uncertainties to the adjusted curves for an arbitrary value of x. However, since for $x = 0$ we have simply $y(0) = A$, it happens that for determining the activity N_o these correlations are irrelevant and that the uncertainty of N_o is therefore just given by $s(N_o) = s(A)$.

In addition, for each fit the program prints out the value of χ^2 and the corresponding number of degrees of freedom, which is given by the denominator of (8).

This computer program, written in FORTRAN IV, is suitable for simultaneous adjustments of up to four branches and it prints out the results for first to third order polynomials in x. The listing is available upon request.

3. Application to activity measurements

The method of simultaneous fitting as described above has been extensively used in the latest two international comparisons of activity measurements organized by the BIPM and which concern ^{134}Cs and ^{137}Cs. For illustrating the practical usefulness of the approach, we restrict ourselves to the case of ^{134}Cs and to our own data; for all details concerning this comparison the full report [5] should be consulted.

When the measurements belonging to the three gamma gates, which are set around about 600, 800 and 1 400 keV, are analyzed independently by adjusting to them polynomials of the form

$$y_k(x) = A + B_k \cdot x + \ldots ,$$

the results for the intercept $y(0) = A = N_o$ are as given in Table 2.

branch	order of fit	$A \pm s(A)$	
1	1	836.2 ± 0.5	
	2	831.0 ± 0.5	X
	3	832.2 ± 1.1	
2	1	830.3 ± 0.2	X
	2	830.6 ± 0.4	
	3	830.5 ± 1.0	
3	1	839.3 ± 1.6	
	2	833.3 ± 2.4	X
	3	834.3 ± 5.8	

Table 2 - Extrapolated activities $N_o = A$ for the BIPM measurements of ^{134}Cs (in units of Bq/mg).

On the basis of the quality of the fit, as indicated by χ^2, it seems justified to choose the values marked by a cross, which is the choice made in [5], but one might also have taken for all branches the result of the second-order fit. The mutual agreement of these extrapolations is satisfactory, although the highest gate (k = 3) adds little.

The evaluation of the weighted mean is done as usually for K results x_k by means of the formula

$$\bar{x} = \frac{\sum_{k=1}^{K} g_k \cdot x_k}{\sum_{k=1}^{K} g_k} \quad , \tag{10}$$

with the statistical weights taken as $g_k = 1/s^2(A)$.

The corresponding uncertainty $S \equiv s_{\bar{x}}$ of the mean can be obtained in two different ways, namely on the basis of "internal" or "external" consistency. The respective formulae for the variances are

- "internal" : $$S^2_{int} = \left(\sum_{k=1}^{K} g_k\right)^{-1} , \tag{11}$$

- "external" : $$S^2_{ext} = \frac{\sum_{k=1}^{K} g_k (x_k - \bar{x})^2}{(K-1)\sum_{k=1}^{K} g_k} . \tag{12}$$

For the extrapolated values x_k (given in Table 2 as A) this leads to a certain choice for the possible "final" result, as indicated in Table 3.

data used	K	$\bar{A}$	$s(\bar{A})$ internal	$s(\bar{A})$ external
same as in [5]	3	830.46 X	0.21 X	0.26
	2	830.44	0.21	0.27
second-order fits	3	830.83	0.33	0.29
	2	830.78	0.33	0.23

Table 3 - Numerical "best values" of $N_o = \bar{A}$, based on the individual fits of branches 1 to 3 (K = 3) or of the branches 1 and 2 (K = 2), with their statistical uncertainties $s(\bar{A})$. The units are the same as in Table 2.

There may be some hesitation about the "best method" to adopt (number of branches, internal or external uncertainty); in our example it happens fortunately that the choice is not a very critical one as there is good consistency between the various results given in Table 3.

Application of the method of simultaneous fitting leads to quite comparable results, but it has at least the advantage that no choice is needed for the uncertainties. For a second-order fit one obtains

- for branches 1 to 3 : $A = 831.04 \pm 0.41$,
- " " 1 and 2 : $A = 830.83 \pm 0.34$.

The least-squares adjustment of the actual measurements to all three branches is shown in Fig. 2.

As we see no compelling reason to eliminate the third decay branch, although it apparently carries little information, the final BIPM result for this intercomparison, obtained with simultaneous fitting, might be stated as

$$N_o = (831.0 \pm 0.4)\ \text{Bq/mg}\ ,$$

which is in good agreement with [5]. The now somewhat larger uncertainty is in line with the data of Table 2 and may be welcome.

4. Final remarks

The method described above permits a simultaneous fitting of several decay branches measured for standardizing a radionuclide where several gamma transitions can be observed.

Although some of the problems raised by the need to combine somehow the usual individual extrapolations are now eliminated as the pooling is accomplished automatically and in a natural way by this method, it should be kept in mind that there are also clear limits for its reasonable application. In particular, if the individual extrapolations to $x = 0$ lead to manifestly incompatible values, a simultaneous fit cannot lead to reliable results as the method obviously assumes a set of coherent data. In this case, it will be necessary to find out and eliminate the causes for the bias before a meaningful adjustment can be made (by any method). If this is not possible, the whole branch leading to an aberrant value of N_o may have to be eliminated and the fit will then be restricted to the remaining curves, but such a decision would clearly need good reasons in order to avoid arbitrariness.

Throughout this exercise it has been assumed that the data taken for the various branches are independent. This would obviously not be true for the experimental techniques permitting to measure several branches at the same time. In this case the more sophisticated technique of multidimensional extrapolation (see [6] to [8]) will be appropriate which accounts for the interdependence of the measured results.

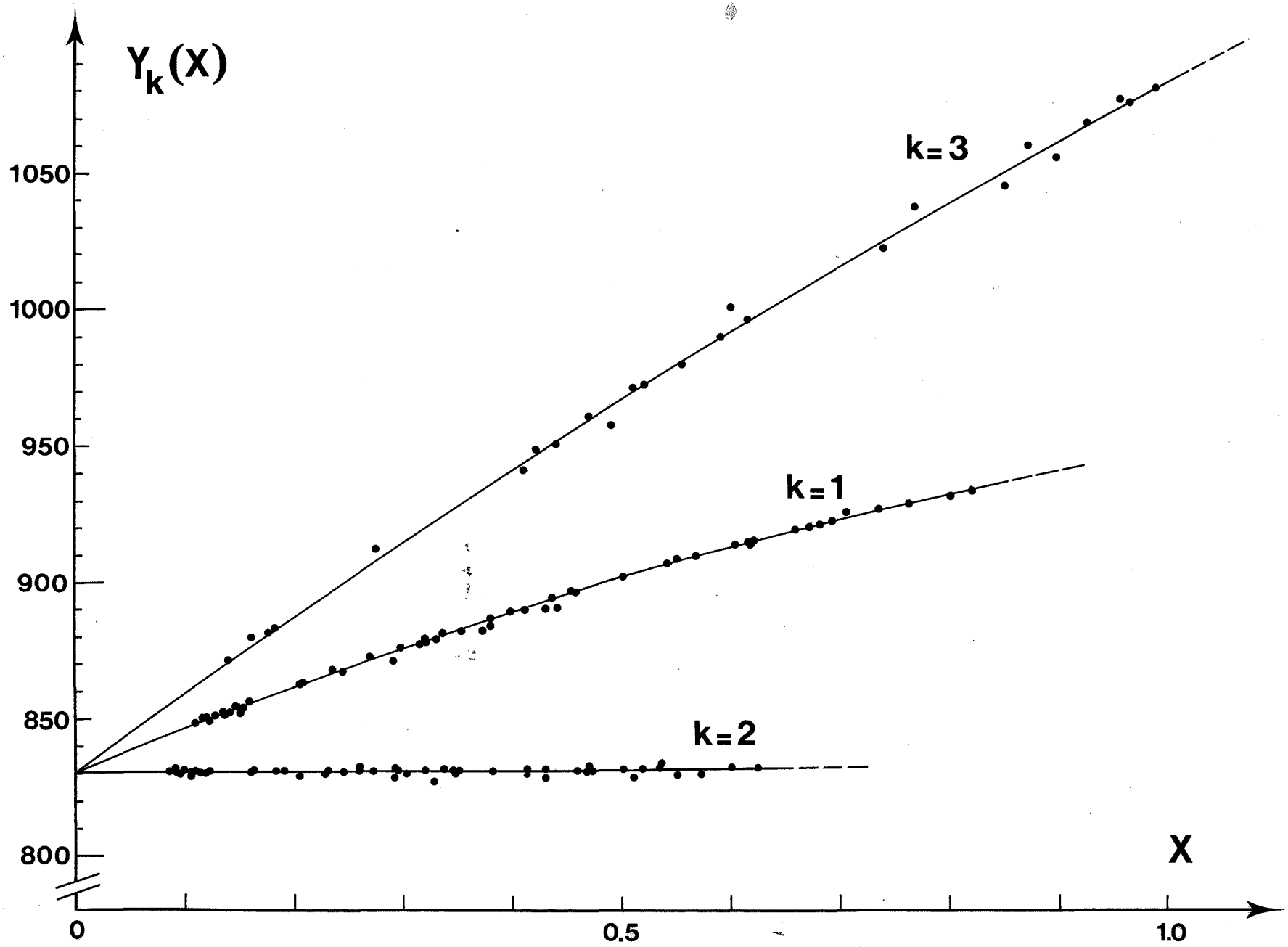

Figure 2 - Second-order simultaneous fit of three decay branches to the measurements of the BIPM performed in connection with the intercomparison of ^{134}Cs. The values of y are in units of Bq/mg for the appropriate reference date (see [5]).

References

[1] "A Handbook of Radioactivity Measurements Procedures", NCRP Report No. 58 (NCRP, Washington, 1978), section 3.2

[2] A. Rytz: "Report on the international comparison of activity measurements of a solution of ^{139}Ce (March 1976)", Rapport BIPM-77/4, p. 17

[3] B.L. van der Waerden: "Mathematical Statistics" (Springer, Berlin, 1969), chapter 7

[4] A. Allisy: "Traitement de données expérimentales", enseignement de métrologie, fascicule 3 (CNAM, Paris, 1977)

[5] A. Rytz: "^{134}Cs, Report of an international comparison of activity measurements (October 1978)", Rapport BIPM-80/2 (to be published)

[6] D. Smith, L.E.H. Stuart: "An extension of the $4\pi\,\beta$-γ coincidence technique: Two dimensional extrapolation", Metrologia 11, 67 (1975)

[7] D. Smith: "An improved method of data collection for 4π β-γ coincidence measurements", Metrologia 11, 73 (1975)

[8] J. Bouchard, R. Vatin: "Développement de la technique de coïncidences 4π β-γ ", Bulletin d'information du BNM, no. 30, p. 30 (1977)

(September 1979)