ornl

ORNL/TM-12106

OAK RIDGE
NATIONAL
LABORATORY

MARTIN MARIETTA

The Covariance Matrix of Derived Quantities and Their Combination

Zhixiang Zhao
F. G. Perey

June 1992

MANAGED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

ORNL/TM-12106

Engineering Physics and Mathematics Division

The Covariance Matrix of Derived Quantities and their Combination

Zhixiang Zhao* and F. G. Perey

*Chinese Nuclear Data Center
P. O. Box 275 (41)
Beijing, P. R. China

DATE PUBLISHED -- June 1992

Prepared for the
Office of Energy Research
Division of Nuclear Physics

Prepared by the
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831
managed by
MARTIN MARIETTA ENERGY SYSTEMS, INC.
for the
U. S. DEPARTMENT OF ENERGY
under contract DE-AC05-84OR21400

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

One of the authors, Zhixiang Zhao, is grateful to R. W. Peelle for constant encouragement, advice, and helpful discussions. He wishes to thank D. C. Larson for his support in this work. He thanks the International Atomic Energy Agency for granting him a fellowship during the course of this work.

The authors are grateful to many individuals, among them R. W. Peelle, F. H. Froehner, S. Chiba, H. Vonach, D. L. Smith, and D. W. Muir, with whom the authors have engaged in fruitful discussions on this subject.

ABSTRACT

The covariance matrix of quantities derived from measured data via nonlinear relations are only approximate since they are functions of the measured data taken as estimates for the true values of the measured quantities. The evaluation of such derived quantities entails new estimates for the true values of the measured quantities and consequently implies a modification of the covariance matrix of the derived quantities that was used in the evaluation process. Failure to recognize such an implication can lead to inconsistencies between the results of different evaluation strategies. In this report we show that an iterative procedure can eliminate such inconsistencies.

1.0 INTRODUCTION

Most data, such as nuclear cross sections, that are reported as having been measured were, in fact, not directly measured. Instead, their values were derived using some quantities that were directly observed in other experiments. Details of how precisely this was done in a specific experiment is often detailed under the heading of "data reduction." Whenever this occurs, the so-called "Law of Error Propagation (LEP)" is used to generate the covariance matrix of these derived quantities from the uncertainties associated with the directly measured quantities. Even though this LEP is well known to all experimenters and evaluators, there appears to be some circumstances where differences of opinion seem to exist today concerning its proper application. We became aware of this through informal discussions of a problem proposed by R. W. Peelle (1) which has been referred to as "Peelle's Pertinent Puzzle (PPP)."

In essence, what has been referred to as PPP is based upon the following situation: we are given two independent measurements of a physical quantity A, which we will denote by A1 and A2, and another independent measurement of a different physical quantity C, which we will denote C1. We are also given the standard deviations associated with these measurements, i.e. the so called "experimental errors". Let us now suppose that we are interested in the physical quantity which is given by the ratio A/C. The most straight forward way of estimating the value for this ratio is to first combine, using the method of least squares, the two independent measurements for A, and then this least squares estimate for A is divided by the measured value for C. The uncertainty in this ratio is obtained by "propagating the errors" in the usual fashion, i.e. by using the LEP. There is an alternate way of proceeding which uses also the method of least squares and the LEP, but this second method involves generating a non-diagonal covariance matrix. From the three independent measurements one derives two values for the ratio of interest: A1/C1 and A2/C1. These two derived values for the ratio of interest are not independent of each other since the same measurement of C was used in deriving them. Consequently, the covariance matrix associated with these two derived values is not diagonal. We can nevertheless still use the method of least squares to combine these two derived values in order to obtain an estimate for the ratio of interest. Peelle thought that since the above two ways of proceeding used the same least squares method and LEP, they should both yield the same final answer. However, he was puzzled by the fact that in a particular numerical example, where the nondiagonal covariance matrix associated with the two derived values for the ratio of interest was generated in the usual manner, the two different ways of proceeding yielded different answers. We will show that "Peelle's Pertinent Puzzle" arose from an incorrect application of the Law of Error Propagation, where two different estimates for the true value of A were used in computing the approximate covariance matrix associated with the two derived ratios.

Inconsistencies of the PPP type could readily occur in nuclear data evaluations because most of the input data to these evaluations are derived from directly measured quantities via nonlinear relations. Consequently, the approximate covariance matrix associated with the input data depends implicitly upon the estimates that one is seeking in the evaluation process, and this dependence is usually ignored. In this report we analyze the general problem of generating the covariance matrix associated with derived quantities and the use of such matrices in data evaluations using the least squares method.

In Section 2, we discuss the Law of Error Propagation in a context where a nondiagonal covariance matrix would be generated for two derived quantities. In Section 3, we analyze the general problem of obtaining the least squares estimate of quantities derived from measurements. In Section 4, we discuss the problem of obtaining least squares fits to derived quantities. Numerical examples are provided in Section 5, including the original problem Peelle was considering.

2.0 THE LAW OF ERROR PROPAGATION.

Let us denote by A, B, and C the true values of three physical quantities. We assume that these true values are not known. Let us denote by a_1, b_1, and c_1 three independent measurements of these physical quantities with respective variances $Var(a_1)$, $Var(b_1)$ and $Var(c_1)$. For convenience we introduce a vector notation where vectors and matrices are indicated by a small arrow over the character:

$$\vec{D} \equiv \begin{pmatrix} A \\ B \\ C \end{pmatrix} \quad ; \quad \vec{d}_1 \equiv \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} , \tag{2.1}$$

$$\vec{V}_{d_1} \equiv \begin{pmatrix} Var(a_1) & 0 & 0 \\ 0 & Var(b_1) & 0 \\ 0 & 0 & Var(c_1) \end{pmatrix} \tag{2.2}$$

We further assume that the measurements a_1, b_1, and c_1 are "unbiased and normally distributed." That is to say, we assume that if one were to repeat these measurements one would find the measured values distributed according to the following joint normal density function:

$$P(\vec{d}_1) = \frac{1}{(2\pi)^{3/2} |\vec{V}_{d_1}|^{1/2}} \ \exp \ (\, - 1/2 \, Q \, (\vec{d}_1) \,) \tag{2.3}$$

$$Q(\vec{d}_1) = (\vec{d}_1 - \vec{D})^t \vec{V}_{d_1}^{-1} (\vec{d}_1 - \vec{D}) \tag{2.4}$$

Alternatively, we assume that one can consider that the values denoted a_1, b_1, and c_1 were randomly selected from such a density function. We denote by $\delta\vec{d}_1$ the deviations of the measurements $\vec{d}_1$ from the true values $\vec{D}$:

$$\delta\vec{d}_1 \equiv (\vec{d}_1 - \vec{D}) \tag{2.5}$$

and, using angle brackets to denote expectation values, we formally have:

$$\vec{V}_{d_1} = \langle \delta\vec{d}_1 \cdot \delta\vec{d}_1^{\,t} \rangle \tag{2.6}$$

We now introduce a physical quantity, whose true value we denote X, that is related to the physical quantities whose true values are denoted A, B and C as follows:

$$\begin{aligned} X &= F(A,C) \\ X &= f(B,C) \end{aligned} \tag{2.7}$$

Consequently, from the measurements a_1, b_1 and c_1 we can derive two different estimated values for X:

$$\begin{aligned} x_1 &= F\,(a_1, c_1) \\ x_2 &= f\,(b_1, c_1) \end{aligned} \tag{2.8}$$

The functions F and f need not be different. When these functions are different we are then dealing with three different physical quantities whose true values we denoted A, B and C. However, when we say that these functions are the same, then the quantities whose true values were denoted A and B are not different physical quantities and the measurements denoted by a_1 and b_1 are to be interpreted as two different and independent measurements of the same physical quantity whose true value is denoted A.

Let us now introduce a vector $\vec{g}_1$ whose components are x_1, x_2 and c_1, i.e.

$$\vec{g}_1 \equiv \begin{pmatrix} x_1 \\ x_2 \\ c \end{pmatrix} \tag{2.9}$$

and we refer to $\vec{g}_1$ as the derived data vector. The so called Law of Error Propagation deals with generating the covariance matrix $\vec{V}_{g_1}$ that we should associate with $\vec{g}_1$.

2.1 LINEAR FUNCTIONS.

Let us assume that the functions $F(A, C)$ and $f(B,C)$ are linear functions:

$$\begin{aligned} X &= F(A,C) = \lambda_0 + \lambda_1 A + \lambda_2 C \\ X &= f(B,C) = \alpha_0 + \alpha_1 B + \alpha_2 C \end{aligned} \tag{2.10}$$

Then we have:

$$\begin{aligned} x_1 &= \lambda_0 + \lambda_1 a_1 + \lambda_2 c_1 \\ x_2 &= \alpha_0 + \alpha_1 b_1 + \alpha_2 c_1 \end{aligned} \tag{2.11}$$

The deviations of the derived quantities are:

$$\begin{aligned} \delta x_1 &= x_1 - X \\ &= \lambda_1(a_1 - A) + \lambda_2(c_1 - C) \\ &= \lambda_1 \delta a_1 + \lambda_2 \delta c_1 \end{aligned} \tag{2.12}$$

and

$$\delta x_2 = \alpha_1 \delta b_1 + \alpha_2 \delta c_1 \tag{2.13}$$

The elements of the covariance matrix $\vec{V}_{g_1}$ are readily computed; for instance, we have:

$$
\begin{aligned}
Var\ (x_1) &= \langle \delta x_1^2 \rangle, \\
&= \int (\lambda_1 \delta a_1 + \lambda_2 \delta c_1)^2\, P(\vec{d}_1)\, d\vec{d}_1, \\
&= \lambda_1^2\ Var(a_1) + \lambda_2^2\ Var(c_1)
\end{aligned}
\tag{2.14}
$$

and similarly:

$$
\begin{aligned}
Var\ (x_2) &= \alpha_1^2\ Var(b_1) + \alpha_2^2\ Var(c_1) \\
Cov\ (x_1, x_2) &= \alpha_2 \lambda_2\ Var(c_1) \\
Cov\ (x_1, c_1) &= \lambda_2\ Var(c_1) \\
Cov\ (x_2, c_1) &= \alpha_2\ Var(c_1)
\end{aligned}
\tag{2.15}
$$

In matrix notation we have:

$$
\vec{G} = \begin{pmatrix} X \\ X \\ C \end{pmatrix}
\tag{2.16}
$$

The deviation $\delta\vec{g}_1$ of the derived data vector $\vec{g}_1$ is:

$$
\delta\vec{g}_1 = \vec{g}_1 - \vec{G} = \vec{S}\ \delta\vec{d}_1
\tag{2.17}
$$

with:

$$
\vec{S} = \begin{pmatrix} \lambda_1 & 0 & \lambda_2 \\ 0 & \alpha_1 & \alpha_2 \\ 0 & 0 & 1 \end{pmatrix}
\tag{2.18}
$$

The covariance matrix $\vec{V}_{g_1}$ associated with the derived data vector $\vec{g}_1$ becomes:

$$\begin{aligned}\vec{V}_{g_1} &= <\delta\vec{g}_1 \cdot \delta\vec{g}_1^t> \\ &= \vec{S} <\delta\vec{d}_1 \cdot \delta\ \vec{d}_1^t> \vec{S}^t \\ &= \vec{S}\ \vec{V}_{d_1}\ \vec{S}^t\end{aligned} \tag{2.19}$$

The quadratic form (2.4) can also be readily transformed:

$$\delta\vec{d}_1^t \vec{V}_{d_1}^{-1} \delta\vec{d}_1 = \delta\vec{g}_1^t \vec{V}_{g_1}^{-1} \delta\vec{g}_1 \tag{2.20}$$

2.2 NONLINEAR FUNCTIONS.

When the functions $F(A, C)$ and $f(B,C)$ are nonlinear, truncated Taylor expansions about the true values A, B and C are used to linearize them. With such truncated expansions the deviations of the derived values are linear functions of the deviations of the measured values and one can then readily calculate the covariance matrix of the derived values. Proceeding formally we have:

$$x_1 = \sum_{m=0}^{\infty} \frac{1}{m!}\left[(a_1 - A)\frac{\partial}{\partial a} + (c_1 - c)\frac{\partial}{\partial c}\right]^m F(a,c)\bigg|_{a=A,\ c=C} \tag{2.21}$$

$$\approx F(A,C) + F_a \delta a_1 + F_c \delta c_1 \tag{2.22}$$

and

$$x_2 = \sum_{m=0}^{\infty} \frac{1}{m!}\left[(b_1 - B)\frac{\partial}{\partial b} + (c_1 - C)\frac{\partial}{\partial c}\right]^m f(b,c)\bigg|_{b = B,\ c = C} \tag{2.23}$$

$$\approx f(B,C) + f_b \delta b_1 + f_c \delta c_1 \tag{2.24}$$

where:

$$F_a \equiv \left.\frac{\partial F(a,c)}{\partial a}\right|_{\substack{a=A\\c=C}} \quad ; \quad F_c \equiv \left.\frac{\partial F(a,c)}{\partial c}\right|_{\substack{a=A\\c=C}} \quad ;$$

$$f_b \equiv \left.\frac{\partial f(b,c)}{\partial b}\right|_{\substack{b=B\\c=C}} \quad ; \quad f_c \equiv \left.\frac{\partial f(b,c)}{\partial c}\right|_{\substack{b=B\\c=C}} \tag{2.25}$$

Using (2.7) we obtain from (2.22) and (2.24):

$$\begin{aligned} \delta x_1 &= F_a \delta a_1 + F_c \delta c_1 \quad , \\ \delta x_2 &= f_b \delta b_1 + f_c \delta c_1 \quad . \end{aligned} \tag{2.26}$$

In practice, the true values A, B and C are not known and the terms in the expansions (2.22) and (2.24) must be calculated at some estimate $\hat{A}$, $\hat{B}$ and $\hat{C}$ for the true values A, B and C. Consequently, when the derived quantities are nonlinear functions of the measured quantities, we can only obtain an estimate of the linear dependence of the deviations of the derived values upon the deviations of the measured values. In order to remind us that this is the case, we introduce the notation:

$$\hat{F}_a \equiv \left.\frac{\partial F(a,c)}{\partial a}\right|_{\substack{a=\hat{A}\\c=\hat{C}}} \quad ; \quad \hat{F}_c \equiv \left.\frac{\partial F(a,c)}{\partial c}\right|_{\substack{a=\hat{A}\\c=\hat{C}}} ,$$

$$\hat{f}_b \equiv \left.\frac{\partial f(b,c)}{\partial b}\right|_{\substack{b=\hat{B}\\c=\hat{C}}} \quad ; \quad \hat{f}_c \equiv \left.\frac{\partial f(b,c)}{\partial c}\right|_{\substack{b=\hat{B}\\c=\hat{C}}} \tag{2.27}$$

which yields:

$$\begin{aligned} \delta x_1 &\approx \hat{F}_a \delta a_1 + \hat{F}_c \delta c_1 , \\ \delta x_2 &\approx \hat{f}_b \delta b_1 + \hat{f}_c \delta c_1 \quad . \end{aligned} \tag{2.28}$$

We are now in a position to obtain an estimate for the elements of the covariance matrix that we should associate with the derived values. From (2.28), taking expectations over the probability density function (2.3), we obtain:

$$\begin{aligned} \hat{V}ar(x_1) &= \langle \delta x_1^2 \rangle \approx \hat{F}_a^{\,2}\, Var(a_1) + \hat{F}_c^{\,2}\, Var(c_1) , \\ \hat{V}ar(x_2) &= \langle \delta x_c^2 \rangle \approx \hat{f}_b^{\,2}\, Var(b_1) + \hat{f}_c^{\,2}\, Var(c_1) , \\ \hat{C}ov(x_1,x_2) &= \langle \delta x_1 \delta x_2 \rangle \approx \hat{F}_a \hat{f}_c\, Var(c_1) , \\ \hat{C}ov(x_1,c_1) &= \langle \delta x_1 \delta c_1 \rangle \approx \hat{F}_c Var(c_1) , \\ \hat{C}ov(x_2,c_1) &= \langle \delta x_1 \delta c_1 \rangle \approx \hat{f}_c Var(c_1) \end{aligned} \tag{2.29}$$

In a matrix notation similar to the one used in the linear functional dependence case we have:

$$\delta \vec{g}_1 \approx \hat{\vec{S}}\, \delta \vec{d}_1 , \tag{2.30}$$

with now:

$$\hat{\vec{S}} = \begin{pmatrix} \hat{F}_a & 0 & \hat{F}_c \\ 0 & \hat{f}_b & \hat{f}_c \\ 0 & 0 & 1 \end{pmatrix} \tag{2.31}$$

and

$$\hat{\vec{V}}_{g_1} = \langle \delta \vec{g}_1 \cdot \delta \vec{g}_1^{\,t} \rangle \approx \hat{\vec{S}}\ \vec{V}_{d_1} \hat{\vec{S}}^{\,t} \tag{2.32}$$

2.3 CONCLUSION.

In the case where the derived quantities are linear functions of the measured quantities, one can calculate exactly the covariance matrix that one should associate with the derived values from the covariance matrix associated with the measured values. However, when the derived quantities are nonlinear functions of the measured quantities, one can obtain only an estimate for the covariance matrix associated with the derived quantities, and this estimated covariance matrix is based upon some specific estimates for the true values of the measured quantities. Consequently, some consistency problems may arise if such an estimated covariance matrix is subsequently used in conjunction with other results that are themselves based upon different estimates for the true values of the measured quantities. As we will show in Section 5, this is precisely what gave rise to PPP: two different estimates for the same quantity were used in constructing the covariance matrix associated with the two derived values.

In the remainder of this report we examine a few situations where such consistency problems may arise and how to avoid them in practice.

3.0 LEAST SQUARE ESTIMATE OF DERIVED QUANTITIES.

We now analyze two different ways of obtaining a least square estimate for X and show that these two different methods yield identical results. In the first method we fit directly the measured quantities considering them to be functions of X. In the second method we calculate the derived quantities, and their associated covariance matrix, which are then fitted.

3.1 FROM FITTING THE DIRECTLY MEASURED QUANTITIES.

From the directly measured quantities a_1, b_1 and c_1 we can derive two different values for X:

$$\begin{aligned} x_1 &= F(a_1, c_1) \ , \\ x_2 &= f(b_1, c_1) \ , \end{aligned} \tag{3.1}$$

and what we seek is a least square estimate for X. However, in this first method we do not start from the relations (3.1) directly; instead we invert them and consider a_1 and b_1 to be functions of x_1, x_2 and c_1:

$$\begin{aligned} a_1 &= \phi(x_1, c_1) \ , \\ b_1 &= \varphi(x_2, c_1) \ , \end{aligned} \tag{3.2}$$

and we have:

$$\begin{aligned} A &= \phi(X, C) \\ B &= \varphi(X, C) \end{aligned} \tag{3.3}$$

Therefore, we obtain the least square estimate by minimizing with respect to X and C the following quadratic expression:

with:

$$Q(\vec{d}_1) = \delta\vec{d}_1^t \, \vec{V}_{d_1}^{-1} \delta\vec{d}_1 \quad , \tag{3.4}$$

$$\delta\vec{d}_1 = \begin{pmatrix} a_1 - \phi(X,C) \\ b_1 - \varphi(X,C) \\ c_1 - C \end{pmatrix} \tag{3.5}$$

In general $\phi(X,C)$ and $\varphi(X,C)$ are nonlinear functions of X and C. Therefore, we have a classic case of nonlinear least square that is solved by linearization of the functions $\phi(X,C)$ and $\varphi(X,C)$, via a Taylor expansion, and by iteration.

We formally expand $\phi(X,C)$ and $\varphi(X,C)$ to first order about some estimate for X and C, that we will denote $\hat{X}^{(k)}$ and $\hat{C}^{(k)}$, and we will denote $\hat{X}^{(k+1)}$ and $\hat{C}^{(k+1)}$ the values for X and C that minimize (3.4):

$$\begin{aligned} \phi(X,C) &\simeq \hat{\phi}^{(k)} + \hat{\phi}_x^{(k)}(X-\hat{X}^{(k)}) + \hat{\phi}_c^{(k)}(C-\hat{C}^{(k)}) \\ \varphi(X,C) &\simeq \hat{\varphi}^{(k)} + \hat{\varphi}_x^{(k)}(X-\hat{X}^{(k)}) + \hat{\varphi}_c^{(k)}(C-\hat{C}^{(k)}) \end{aligned} \tag{3.6}$$

where:

$$\begin{aligned} &\hat{\phi}^{(k)} \equiv \phi(\hat{X}^{(k)}, \hat{C}^{(k)}) \quad ; \quad \hat{\varphi}^{(k)} \equiv \varphi(\hat{X}^{(k)}, \hat{C}^{(k)}) \quad ; \\ &\hat{\phi}_x^{(k)} \equiv \left.\frac{\partial\phi(x,c)}{\partial x}\right|_{\substack{x = \hat{X}^{(k)} \\ c = \hat{C}^{(k)}}} \quad ; \quad \hat{\varphi}_x^{(k)} \equiv \left.\frac{\partial\varphi(x,c)}{\partial x}\right|_{\substack{x = \hat{X}^{(k)} \\ c = \hat{C}^{(k)}}} \quad ; \\ &\hat{\phi}_c^{(k)} \equiv \left.\frac{\partial\phi(x,c)}{\partial c}\right|_{\substack{x = \hat{X}^{(k)} \\ c = \hat{C}^{(k)}}} \quad ; \quad \hat{\varphi}_c^{(k)} \equiv \left.\frac{\partial\varphi(x,c)}{\partial c}\right|_{\substack{x = \hat{X}^{(k)} \\ c = \hat{C}^{(k)}}} \quad . \end{aligned} \tag{3.7}$$

Substituting (3.6) into (3.5) we obtain:
with:

$$\delta\vec{d}_1 = \vec{d}_1 - \vec{\hat{D}}^{(k)} - \vec{\hat{U}}^{(k)} \left(\vec{P} - \vec{\hat{P}}^{(k)}\right) \tag{3.8}$$

$$\vec{\hat{D}}^{(k)} \equiv \begin{pmatrix} \hat{\Phi}^{(k)} \\ \hat{\varphi}^{(k)} \\ \hat{C}^{(k)} \end{pmatrix} \tag{3.9}$$

$$\vec{\hat{U}}^{(k)} \equiv \begin{pmatrix} \hat{\Phi}_x^{(k)} & \hat{\Phi}_c^{(k)} \\ \hat{\varphi}_x^{(k)} & \hat{\varphi}_c^{(k)} \\ 0 & 1 \end{pmatrix} \tag{3.10}$$

$$\vec{P} \equiv \begin{pmatrix} X \\ C \end{pmatrix} \tag{3.11}$$

$$\vec{\hat{P}}^{(k)} \equiv \begin{pmatrix} \hat{X}^{(k)} \\ \hat{C}^{(k)} \end{pmatrix} \tag{3.12}$$

The quadratic expression (3.4) which is to be minimized with respect to $\vec{P}$ becomes:

$$Q(\vec{d}_1) = \left(\vec{d}_1 - \vec{\hat{D}}^{(k)} - \vec{\hat{U}}^{(k)}(\vec{P} - \vec{\hat{P}}^{(k)})\right)^t \vec{V}_{d_1}^{-1} \left(\vec{d}_1 - \vec{\hat{D}}^{(k)} - \vec{\hat{U}}^{(k)}(\vec{P} - \vec{\hat{P}}^{(k)})\right) \tag{3.13}$$

Denoting by $\vec{\hat{P}}^{(k+1)}$ the value of $\vec{P}$ which minimizes (3.13) and by $\vec{\hat{V}}^{(k+1)}$ the covariance matrix we have:

$$\vec{\hat{P}}^{(k+1)} = \vec{\hat{P}}^{(k)} + \left(\vec{\hat{U}}^{(k)t}\vec{V}_{d_1}^{-1}\vec{\hat{U}}^{(k)}\right)^{-1}\vec{\hat{U}}^{(k)t}\vec{V}_{d_1}^{-1}\left(\vec{d}_1 - \vec{\hat{D}}^{(k)}\right)$$
$$\vec{\hat{V}}^{(k+1)} = \left(\vec{\hat{U}}^{(k)t}\ \vec{V}_{d_1}^{-1}\ \vec{\hat{U}}^{(k)}\right)^{-1} \qquad (3.14)$$

Note: An interesting case is the one where $F(A, C)$ and $F(B, C)$ are the same function. That is to say, what we had denoted b_1 is a second measurement for A. Therefore, in this case we denote b_1 by a_2 and we have:

$$\begin{aligned} x_1 &= F(a_1, c_1) \\ x_2 &= F(a_2, c_1) \end{aligned} \qquad (3.15)$$

The function that we had denoted $\varphi\,(X,C)$ is now identical to $\phi\,(X,C)$. After some straight forward algebra (3.14) yields:

$$\begin{aligned} \hat{X}^{(k+1)} &= \hat{X}^{(k)} + (a_w - \hat{\phi}^{(k)})/\hat{\phi}_x^{(k)} - \hat{\phi}_c^{(k)}(c_1 - \hat{C}^{(k)})/\hat{\phi}_x^{(k)} \\ \hat{C}^{(k+1)} &= c_1 \end{aligned} \qquad (3.16)$$

where a_w is the weighted average of a_1 and a_2:

$$a_w = \frac{a_1/Var(a_1) + a_2/Var(a_2)}{1/Var(a_1) + 1/Var(a_2)} \qquad (3.17)$$

and

$$\hat{X}^{(k)} = F(\hat{\phi}^{(k)}, \hat{C}^{(k)}) \qquad (3.18)$$

Clearly from (3.16) convergence will be obtained when $\hat{C}^{(k)} = c_1$ and $\hat{\phi}^{(k)} = a_w$. Consequently the iteration stops when:

$$\begin{aligned} \hat{X} &= F(a_w, c_1) \\ \hat{C} &= c_1 \end{aligned} \qquad (3.19)$$

What (3.19) states is that the least square estimate for X is obtained from the measurement c_1 and the weighted average of the two measurements a_1 and a_2, a result that could have been anticipated without performing all the algebra.

3.2 FROM FITTING THE DERIVED QUANTITIES.

An alternate way of obtaining a least square estimate for X and for C is to minimize, also with respect to X and to C, a quadratic expression in which the derived values for X, x_1 and x_2, appear explicitly. This quadratic expression is:

$$Q(\vec{g}_1) = \delta\vec{g}_1' \, \hat{\vec{V}}_{g_1}^{-1} \, \delta\vec{g}_1 \quad , \tag{3.20}$$

where:

$$\delta\vec{g}_1 = \begin{pmatrix} x_1 - X \\ x_2 - X \\ c_1 - C \end{pmatrix} = \vec{g}_1 - \vec{T}\vec{P} \quad , \tag{3.21}$$

with:

$$\vec{T} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \tag{3.22}$$

and $\hat{\vec{V}}_{g_1}$ is given by (2.32).

Since $\delta\vec{g}_1$ is a linear function of $\vec{P}$, the elements of $\vec{T}$ being constants, it would seem that we have transformed what was previously a nonlinear least square problem into a linear one. Of course, we could not have done such a thing. As discussed in section 2.2, in order to compute $\hat{\vec{V}}_{g_1}$ we needed some estimate for A, B and C. These estimates are related to and should be consistent with the estimates for X and C that we seek by minimizing (3.20). The estimates $\hat{A}$ and $\hat{B}$ that we should use to calculate $\hat{\vec{V}}_{g_1}$ are related to the solutions $\hat{X}$ and $\hat{C}$ that minimize (3.20) by:

$$\begin{aligned} \hat{A} &= \phi(\hat{X}, \hat{C}) \ , \\ \hat{B} &= \varphi(\hat{X}, \hat{C}) \ , \end{aligned} \tag{3.23}$$

or we should have:

$$\hat{X} = F(\hat{A}, \hat{C}) = f(\hat{B}, \hat{C}) \ . \tag{3.24}$$

The consistency between the values of $\hat{X}$ and $\hat{C}$, that we seek by minimizing (3.20), and the values of $\hat{A}$, $\hat{B}$ and $\hat{C}$, needed to calculate $\hat{\vec{V}}_{g_1}$, that appears in (3.20), can be achieved by iterations. From the k-th estimates $\hat{X}^{(k)}$ and $\hat{C}^{(k)}$ we obtain $\hat{A}^{(k)}$ and $\hat{B}^{(k)}$ via:

$$\hat{A}^{(k)} = \phi\left(\hat{X}^{(k)}, \hat{C}^{(k)}\right) = \hat{\phi}^{(k)} ,$$
$$\hat{B}^{(k)} = \varphi\left(\hat{X}^{(k)}, \hat{C}^{(k)}\right) = \hat{\varphi}^{(k)} . \quad (3.25)$$

Then we use these values of $\hat{A}^{(k)}$, $\hat{B}^{(k)}$ and $\hat{C}^{(k)}$ to calculate the elements of the matrix $\hat{\vec{S}}^{(k)}$:

$$\hat{\vec{S}}^{(k)} = \begin{pmatrix} \hat{F}_a^{(k)} & 0 & \hat{F}_c^{(k)} \\ 0 & \hat{f}_b^{(k)} & \hat{f}_c^{(k)} \\ 0 & 0 & 1 \end{pmatrix} \quad . \quad (3.26)$$

Having obtained such a matrix we calculate $\vec{V}_{g_1}^{(k)}$:

$$\hat{\vec{V}}_{g_1}^{(k)} = \hat{\vec{S}}^{(k)} \vec{V}_{d_1} \hat{\vec{S}}^{(k)t} \quad . \quad (3.27)$$

Furthermore, in order to be fully consistent, the components of $\vec{g}_1$ that appear in (3.21) must be calculated using the same linear approximation that is used to calculate the elements of the covariance matrix $\hat{\vec{V}}_{g_1}^{(k)}$. We will be reminded that this is the case by introducing the notation $\vec{g}_1^{(k)}$ for $\vec{g}_1$ where:

$$\vec{g}_1^{(k)} = \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ c_1^{(k)} \end{pmatrix} \quad (3.28)$$

and

$$
\begin{aligned}
x_1^{(k)} &= \hat{F}^{(k)} + \hat{F}_a^{(k)}(a_1 - \hat{\phi}^{(k)}) + \hat{F}_c^{(k)}(c_1 - \hat{C}^{(k)}) \\
x_2^{(k)} &= \hat{f}^{(k)} + \hat{f}_b^{(k)}(b_1 - \hat{\varphi}^{(k)}) + \hat{f}_c^{(k)}(c_1 - \hat{C}^{(k)}) \\
\hat{C}^{(k)} &= c_1
\end{aligned}
\tag{3.29}
$$

Then we minimize with respect to $\vec{P}$ the quadratic expression:

$$
\left(\vec{g}_1^{(k)} - \vec{T}\vec{P}\right)^t \hat{\vec{V}}_{g_1}^{(k)-1} \left(\vec{g}_1^{(k)} - \vec{T}\vec{P}\right)
\tag{3.30}
$$

If we denote by $\vec{P}^{(k+1)}$ the value of $\vec{P}$ that minimizes (3.30) we have:

$$
\begin{aligned}
\hat{\vec{P}}^{(k+1)} &= \left(\vec{T}^t \hat{\vec{V}}_{g_1}^{(k)-1} \vec{T}\right)^{-1} \vec{T}^t \hat{\vec{V}}_{g_1}^{(k)-1} \vec{g}_1^{(k)} \quad , \\
\hat{\vec{V}}^{(k+1)} &= \left(\vec{T}^t \hat{\vec{V}}_{g_1}^{(k)-1} \vec{T}\right)^{-1} \quad .
\end{aligned}
\tag{3.31}
$$

Let us again consider the special case where b_1 is a second measurement for A that we denote a_2. After some algebra (3.31) yields:

$$
\begin{aligned}
\hat{X}^{(k+1)} &= F(\hat{A}^{(k)}, \hat{C}^{(k)}) + \hat{F}_a^{(k)}(a_w - \hat{A}^{(k)}) + \hat{F}_c^{(k)}(c_1 - \hat{C}^{(k)}) \\
\hat{C}^{(k+1)} &= c_1
\end{aligned}
\tag{3.32}
$$

where a_w is given by (3.17) as before.

This solution is identical to the one found before, (3.16), since:

$$
\begin{aligned}
\hat{A}^{(k)} &= \hat{\phi}^{(k)} \quad , \\
\hat{F}_a^{(k)} &= 1 / \hat{\phi}_x^{(k)} \quad , \\
\hat{F}_c^{(k)} &= -\hat{\phi}_c^{(k)} / \hat{\phi}_x^{(k)} \quad .
\end{aligned}
\tag{3.33}
$$

and again the iterations in (3.32) will stop when:

$$\hat{X} = F(a_w, c_1) \qquad \hat{C} = c_1 \tag{3.34}$$

As we now show, this is not a special case; the solutions (3.14) and (3.31) will always yield identical results whatever the functions $F(A,\ C)$ and $f(A,\ C)$ are.

In (3.31) we replace $\vec{\hat{V}}_{g_1}^{(k)}$ by its expression (3.27) to obtain:

$$\vec{\hat{P}}^{(k+1)} = \left(\vec{T}^t\, \vec{\hat{S}}^{(k)-t}\, \vec{V}_{d_1}^{-1}\, \vec{\hat{S}}^{(k)-1}\, \vec{T}\right)^{-1} \vec{T}^t\, \vec{\hat{S}}^{(k)-t}\, \vec{V}_{d_1}^{-1}\, \vec{\hat{S}}^{(k)-1}\, \vec{g}_1^{(k)} \qquad ; \tag{3.35}$$

but we have:

$$\begin{aligned}\vec{\hat{S}}^{(k)-1}\vec{T} &= \begin{pmatrix} 1/\hat{F}_a^{(k)} & 0 & -\hat{F}_c^{(k)}/\hat{F}_a^{(k)} \\ 0 & 1/\hat{f}_b^{(k)} & -\hat{f}_c^{(k)}/\hat{f}_b^{(k)} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1/\hat{F}_a^{(k)} & -\hat{F}_c^{(k)}/\hat{F}_a^{(k)} \\ 1/\hat{f}_b^{(k)} & -\hat{f}_c^{(k)}/\hat{f}_b^{(k)} \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \hat{\Phi}_x^{(k)} & \hat{\Phi}_c^{(k)} \\ \hat{\varphi}_x^{(k)} & \hat{\varphi}_c^{(k)} \\ 0 & 1 \end{pmatrix} = \vec{\hat{U}}^{(k)}\end{aligned} \tag{3.36}$$

Replacing $\vec{\hat{S}}^{(k)-1}\vec{T}$ by $\vec{\hat{U}}^{(k)}$ in (3.35) we obtain:

$$\hat{\vec{P}}^{(k+1)} = \left(\hat{\vec{U}}^{(k)t}\vec{V}_{d_1}^{-1}\hat{\vec{U}}^{(k)}\right)^{-1}\hat{\vec{U}}^{(k)t}\vec{V}_{d_1}^{-1}\hat{\vec{S}}^{(k)-1}\vec{g}_1^{(k)} \tag{3.37}$$

$$\hat{\vec{P}}^{(k)} = \left(\hat{\vec{U}}^{(k)t}\vec{V}_{d_1}^{-1}\hat{\vec{U}}^{(k)}\right)^{-1}\left(\hat{\vec{U}}^{(k)t}\vec{V}_{d_1}^{-1}\hat{\vec{U}}^{(k)}\right)\hat{\vec{P}}^{(k)} \tag{3.38}$$

$$\hat{\vec{P}}^{(k+1)} = \hat{\vec{P}}^{(k)} + \left(\hat{\vec{U}}^{(k)t}\vec{V}_{d_1}^{-1}\hat{\vec{U}}^{(k)}\right)^{-1}\hat{\vec{U}}^{(k)t}\vec{V}_{d_1}^{-1}\left(\hat{\vec{S}}^{(k)-1}\vec{g}_1^{(k)} - \hat{\vec{U}}^{(k)}\hat{\vec{P}}^{(k)}\right) \tag{3.39}$$

The expression (3.39) is identical to the result (3.14) after suitable manipulation of the last term on the right hand side as follows:

$$\begin{aligned}
\hat{\vec{S}}^{(k)-1}\vec{g}_1^{(k)} - \hat{\vec{U}}^{(k)}\hat{\vec{P}}^{(k)} &= \hat{\vec{S}}^{(k)-1}\left(\vec{g}_1^{(k)} - \hat{\vec{S}}^{(k)}\hat{\vec{U}}^{(k)}\hat{\vec{P}}^{(k)}\right) \\
&= \hat{\vec{S}}^{(k)-1}\left(\vec{g}_1^{(k)} - \vec{T}\hat{\vec{P}}^{(k)}\right) \\
&= \hat{\vec{S}}^{(k)-1}\left(\vec{g}_1^{(k)} - \vec{T}\vec{P} + \vec{T}(\vec{P} - \hat{\vec{P}}^{(k)})\right) \\
&= \hat{\vec{S}}^{(k)-1}\left(\vec{g}_1^{(k)} - \vec{G}\right) + \hat{\vec{S}}^{(k)-1}\vec{T}\left(\vec{P} - \hat{\vec{P}}^{(k)}\right) \\
&= \vec{d}_1 - \vec{D} - \hat{\vec{U}}^{(k)}\left(\hat{\vec{P}}^{(k)} - \vec{P}\right) \\
&= \vec{d}_1 - \hat{\vec{D}}^{(k)}
\end{aligned} \tag{3.40}$$

Therefore, the two different ways of obtaining the least square estimate of the derived quantities will always produce identical results.

4.0 LEAST SQUARE FIT TO DERIVED QUANTITIES.

We frequently least square fit theoretical expressions to data whose values were not directly measured but were derived from other measured data. If the functional dependence of the data being fitted upon the measured quantities is nonlinear, then, as discussed in section 2, the covariance matrix associated with the data being fitted is only approximate. We now show that an "outer iteration" may be required in the least square fitting process to deal with these non linearities. We also show that such "outer iterations" are not necessary if one fits the directly measured data instead of the derived data. Let us denote by:

$$X(E) = R(E;H) \ , \tag{4.1}$$

the theoretical expression being fitted to the derived data, where E is the independent variable and the L components of the abstract vector H are the parameters being adjusted to fit the data. We consider the situation where the $X(E)$'s are related to some other physical quantities we denote $A(E)$ and C via a nonlinear relation:

$$X(E) = F(A(E\ ), C). \tag{4.2}$$

Let us assume that we have n measurements of the quantities $A(E)$ and we denote these measurements:

$$a_1(E_i) \equiv a_1(i) \ , \ i = 1, n \qquad , \tag{4.3}$$

and consider the $a_1(i)$'s to be the components of an abstract vector $\vec{a}_1$:

$$\vec{a}_1 \equiv \begin{pmatrix} a_1(1) \\ a_1(2) \\ \vdots \\ a_1(n) \end{pmatrix} \tag{4.4}$$

In addition to these n measurements let us denote by c_1 a measurement of the physical quantity denoted C in (4.2). For convenience sake we consider these $n+1$ measurements to be independent measurements and that we can neglect the uncertainties in the variables E_i where the measurements we denote $a_1(i)$ were made. We introduce what we refer to as the "directly measured data" vector that we denote $\vec{d}_1$ with:

$$\vec{d}_1 \equiv \begin{pmatrix} a_1(1) \\ a_1(2) \\ \vdots \\ a_1(n) \\ c_1 \end{pmatrix} = \begin{pmatrix} \vec{a}_1 \\ c_1 \end{pmatrix} \quad , \tag{4.5}$$

and we denote by $\vec{V}_{d_1}$ the diagonal covariance matrix associated with the data vector $\vec{d}_1$.

From $\vec{a}_1$ and c_1, using (4.2), we obtain:

$$x_1(E_i) \equiv x_1(i) \equiv F(a_1(E_i)\ ,\ c_1)\ , \tag{4.6}$$

and consider the $x_1(i)$'s to be the components of an abstract vector $\vec{x}_1$. As we did in section 3, we introduce a "derived data" vector $\vec{g}_1$ whose components are:

$$\vec{g}_1 \equiv \begin{pmatrix} \vec{x}_1 \\ \\ c_1 \end{pmatrix} \tag{4.7}$$

We now need to express the deviation $\delta\vec{g}_1$ in terms of the deviation $\delta\vec{d}_1$ in order to generate the covariance matrix $\vec{V}_{g_1}$ associated with the derived data vector. As discussed in section 2, the elements of the transformation matrix that transforms $\delta\vec{d}_1$ into $\delta\vec{g}_1$ are the partial derivatives of $F(A(E), C)$ evaluated at some estimate of the true values $A(E)$ and C. Let us denote by $\hat{A}^{(k)}_{(i)}$ and $\hat{C}^{(k)}$ such estimates, and we have:

$$\delta\vec{g}_1 = \vec{\vec{S}}^{(k)} \delta\vec{d}_1 \tag{4.8}$$

where:

$$
\hat{S}^{(k)} \equiv \begin{pmatrix}
\hat{F}^{(k)}_{a(1)} & 0 & 0 & \cdot & \cdot & \cdot & 0 & \hat{F}^{(k)}_{c(1)} \\
0 & \hat{F}^{(k)}_{a(2)} & 0 & \cdot & \cdot & \cdot & 0 & \hat{F}^{(k)}_{c(2)} \\
0 & 0 & \hat{F}^{(k)}_{a(3)} & \cdot & \cdot & \cdot & 0 & \hat{F}^{(k)}_{c(3)} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & \cdot & \cdot & \cdot & \hat{F}^{(k)}_{a(n)} & \hat{F}^{(k)}_{c(n)} \\
0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 1
\end{pmatrix} \tag{4.9}
$$

$$
\hat{F}^{(k)}_{a(i)} \equiv \left. \frac{\partial F(a(i),c)}{\partial a(i)} \right|_{\substack{a(i) = \hat{A}^{(k)}_{(i)} \\ c = \hat{C}^{(k)}}}
$$

$$
\hat{F}^{(k)}_{c(i)} \equiv \left. \frac{\partial F(a(i),c)}{\partial c} \right|_{\substack{a(i) = \hat{A}^{(k)}_{(i)} \\ c = \hat{C}^{(k)}}}
$$

The $\hat{A}^{(k)}(i)$ must be consistent with the estimate we will obtain for the true values of the parameters we denote $\vec{H}^{(k)}$. If we denote the inverse of (4.2):

$$
A(E) = \phi(X(E), C) = \phi(R(E;\vec{H}), C) \ , \tag{4.10}
$$

then:

$$
\hat{A}^{(k)}(i) = \hat{A}^{(k)}(E_i) = \phi(R(E_i;\vec{H}^{(k)}), \hat{C}^{(k)}) \quad . \tag{4.11}
$$

From (4.8) we obtain:

$$\hat{\vec{V}}_{g_1}^{(k)} = \hat{\vec{S}}^{(k)} \vec{V}_{d_1} \hat{\vec{S}}^{(k)t} \quad . \tag{4.12}$$

The quadratic form we want to minimize with respect to $\vec{G}$ is:

$$Q(\vec{g}_1) = \left(\vec{g}_1^{(k)} - \vec{G}\right)^t \hat{\vec{V}}_{g_1}^{(k)-1} \left(\vec{g}_1^{(k)} - \vec{G}\right) \quad , \tag{4.13}$$

where the components of $\vec{G}$ are given by:

$$\begin{aligned} \vec{G}(i) &= R(E_i; \vec{H}) \ , \ i \le n \\ \vec{G}(n+1) &= C \end{aligned} \quad . \tag{4.14}$$

Introducing the parameter vector $\vec{P}$, where:

$$\vec{P} = \begin{pmatrix} \vec{H} \\ C \end{pmatrix} , \tag{4.15}$$

we formally express $\vec{G}$ as:

$$\vec{G} = \hat{\vec{G}}^{(k)} + \hat{\vec{T}}^{(k)} (\vec{P} - \hat{\vec{P}}^{(k)}) \tag{4.16}$$

where:

$$\begin{aligned} \hat{\vec{G}}^{(k)}(i) &= R(E_i; \hat{\vec{H}}^{(k)}) , i \le n \\ \hat{\vec{G}}^{(k)}(n+1) &= \hat{C}^{(k)} \end{aligned} \quad , \tag{4.17}$$

and

$$
\begin{aligned}
\vec{\hat{T}}^{(k)}(i,j) &= \left.\frac{\partial R(E_i\,;\vec{h})}{\partial \vec{h}(j)}\right|_{\vec{h} = \vec{\hat{H}}^{(k)}} && \textit{for } i \le n, \textit{ and } j \le L \\
\vec{\hat{T}}^{(k)}(n+1) &= 0 && \textit{for } j \le L \\
\vec{\hat{T}}^{(k)}(i,L+1) &= 0 && \textit{for } i \le n \\
\vec{\hat{T}}^{(k)}(n+1,L+1) &= 1
\end{aligned}
\tag{4.18}
$$

and also

$$
\vec{\hat{P}}^{(k)} = \begin{pmatrix} \vec{\hat{H}}^{(k)} \\ \vec{\hat{C}}^{(k)} \end{pmatrix} . \tag{4.19}
$$

Substituting (4.16) for $\vec{G}$ in (4.13) we obtain:

$$
Q(\vec{g}_1) = \left(\vec{g}_1^{(k)} - \vec{\hat{G}}^{(k)} - \vec{\hat{T}}^{(k)}\,(\vec{P} - \vec{\hat{P}}^{(k)})\right)^t \vec{\hat{V}}_{g_1}^{(k)-1}\left(\vec{g}_1^{(k)} - \vec{\hat{G}}^{(k)} - \vec{\hat{T}}^{(k)}\,(\vec{P} - \vec{\hat{P}}^{(k)})\right) \tag{4.20}
$$

If we denote by $\vec{\hat{P}}^{(k+1)}$ the value of $\vec{P}$ that minimizes (4.20) and by $\vec{\hat{V}}^{(k+1)}$ its associated covariance matrix, we have:

$$
\begin{aligned}
\vec{\hat{P}}^{(k+1)} &= \vec{\hat{P}}^{(k)} + \left(\vec{\hat{T}}^{(k)t}\,\vec{\hat{V}}_{g_1}^{(k)-1}\,\vec{\hat{T}}^{(k)}\right)^{-1}\vec{\hat{T}}^{(k)t}\,\vec{\hat{V}}_{g_1}^{(k)-1}\left(\vec{g}_1^{(k)} - \vec{\hat{G}}^{(k)}\right) \\
\vec{\hat{V}}^{(k+1)} &= \left(\vec{\hat{T}}^{(k)t}\,\vec{\hat{V}}_{g_1}^{(k)-1}\,\vec{\hat{T}}^{(k)}\right)^{-1}
\end{aligned}
\tag{4.21}
$$

It should be noted that if $R(E;\vec{H})$ is a linear function of the L components of $\vec{H}$ then (4.16) is exact, and the notation $\vec{\hat{T}}^{(k)}$ is somewhat misleading since the elements of $\vec{\hat{T}}^{(k)}$ are constants. However, even if $R(E;H)$ is a linear function of the L components of $\vec{H}$, one must iterate (4.24) because we have assumed that the function $F(A(E);C)$ was nonlinear and consequently the covariance matrix of the derived data $\hat{\vec{V}}_{g_1}^{(k)}$ is an approximation which will be a function of $\hat{\vec{H}}^{(k)}$. To our knowledge, very few if any of the least square fitting codes that allow for a non-diagonal covariance matrix of the data being fitted allow for an iterated solution of the type represented by (4.24) where one recalculates the covariance matrix of the data being fitted based upon an updated estimate for the parameters being determined.

However, even though what we seek through the fitting process are the parameters of a theoretical expression for the derived data, it is not necessary that we directly fit the derived data $\vec{g}_1$. The same end result can be achieved by fitting the directly measured data $\vec{d}_1$ with its diagonal covariance matrix $\vec{V}_{d_1}$, in a way similar to what was done in section 2.

Let us denote by $\vec{D}$ the abstract vector whose components are the true values of the components of the directly measured data vector $\vec{d}_1$. We have:

$$\vec{D} = \begin{pmatrix} \phi(R(E_1;\vec{H}),\, C) \\ \phi(R(E_2;\vec{H}),\, C) \\ \vdots \\ \phi(R(E_n;\vec{H}),\, C) \\ C \end{pmatrix} \tag{4.22}$$

Treating $\vec{D}$ as a nonlinear function of $\vec{H}$ and C, the components of $\vec{P}$, we can obtain the least square estimate for $\vec{H}$ that we seek by minimizing with respect to $\vec{P}$ the following quadratic expression:

$$Q(\vec{d}_1) = (\vec{d}_1 - \vec{D})^t\, \vec{V}_{d_1}^{\,-1}(\vec{d}_1 - \vec{D}) \tag{4.23}$$

As usual this nonlinear least square problem can be solved by iteration. Performing a Taylor expansion of $\vec{D}$ about some estimate for $\vec{P}$ we have:

$$\vec{D} = \vec{\hat{D}}^{(k)} + \vec{\hat{U}}^{(k)}(\vec{P} - \vec{\hat{P}}^{(k)}) \quad , \tag{4.24}$$

with:

$$
\begin{aligned}
\vec{\hat{D}}^{(k)}(i) &= \phi\left(R(E_i\,;\,\vec{\hat{H}}^{(k)}),\hat{C}^{(k)}\right) , && \textit{for } i \le n \\
\vec{\hat{D}}^{(k)}(n+1) &= \hat{C} \\
\vec{\hat{U}}^{(k)}(i,j) &= \left.\frac{\partial\phi\left(R(E_i\,;\,\vec{h}\,),c\right)}{\partial\vec{h}\,(j)}\right|_{\substack{\vec{h}\,=\,\vec{\hat{H}}^{(k)} \\ c\,=\,\hat{C}^{(k)}}} , && \textit{for } i \le n \textit{ and } j \le L \\
\vec{\hat{U}}^{(k)}(i,L+1) &= \left.\frac{\partial\phi\left(R(E_i;\vec{h}\,)c\right)}{\partial c}\right|_{\substack{\vec{h}\,=\,\vec{\hat{H}}^{(k)} \\ c\,=\,\hat{C}^{(k)}}} , && \textit{for } i \le n \\
\vec{\hat{U}}^{(k)}(n+1,j) &= 0 && \textit{for } j \le L \\
\vec{\hat{U}}^{(k)}(n+1,L+1) &= 1
\end{aligned}
\tag{4.25}
$$

Substituting the Taylor expansion (4.24) into (4.23) we obtain:

$$
Q\,(\vec{d}_1) = \left(\vec{d}_1 - \vec{\hat{D}}^{(k)} - \vec{\hat{U}}^{(k)}(\vec{P} - \vec{\hat{P}}^{(k)})\right)^t \vec{V}_{d_1}^{-1} \left(\vec{d}_1 - \vec{\hat{D}}^{(k)} - \vec{\hat{U}}^{(k)}(\vec{P} - \vec{\hat{P}}^{(k)})\right) , \tag{4.26}
$$

which is formally identical to the quadratic expression (3.13). Consequently the result sought is formally given by (3.14).

5.0 COMPUTATIONAL EXAMPLES.

In this section several computational examples are given, and we illustrate some of the inconsistencies that can arise when one ignores the fact that covariance matrices of derived data are only approximate if they are not linearly related to the directly measured data.

5.1 LINEAR FUNCTIONS.

In this example we take the functions $F(A,C)$ and $f(B,C)$ to be different but linear:

$$\begin{aligned} X &= F(A,C) = A - C \\ X &= f(B,C) = B - 2C/3 \end{aligned} \tag{5.1}$$

and the independently measured values, with their standard deviations, to be:

$$\begin{aligned} a_1 &= 2.5 \pm 0.15 \\ b_1 &= 1.67 \pm 0.10 \\ c_1 &= 1.0 \pm 0.30 \end{aligned} \tag{5.2}$$

Let us calculate the least square estimated values for A, B, C and X. At first sight it would seem that: (1) since we have single independently measured values for A, B and C their least square estimate should be these measured values, and (2) since using (5.1) we can derive two values for X,

$$\begin{aligned} x_1 &= a_1 - c_1 = 1.5 \quad , \\ x_2 &= b_1 - \frac{2c_1}{3} = 1.0 \quad , \end{aligned} \tag{5.3}$$

the least square estimate for X will be between these two derived values. The situation is in fact slightly more complicated since we can eliminate X from the relations (5.1) to obtain:

$$A = \frac{(C + 3B)}{3} \quad ,$$
$$B = \frac{(3A - C)}{3} \quad , \tag{5.4}$$
$$C = 3(A - B) \quad ,$$
$$X = 3B - 2A$$

Therefore, given the measurements b_1 and c_1 we can derive a value for A which is independent from the directly measured value a_1. The least square estimate for A should be based upon these two "measurements" for A. The same situation prevails for B and C. Consequently we have not only a_1, b_1 and c_1 but also:

$$a_2 = \frac{(c_1 + 3b_1)}{3} = 2.0$$
$$b_2 = \frac{(3a_1 - c_1)}{3} = 2.167 \tag{5.5}$$
$$c_2 = 3(a_1 - b_1) = 2.5$$

Furthermore, from (5.1) we can extract the identity:

$$X = 3B - 2A, \tag{5.6}$$

which allows us to obtain a third derived value for X that is not based upon the measurement c_1, whereas the other two derived values for X were based upon it:

$$x_3 = 3b_1 - 2a_1 = 0.0. \tag{5.7}$$

It should be noted that the three derived values for X are not independent, and it is no longer obvious that the least square estimate should be between the derived values denoted x_1 and x_2.

Let us now apply the formalism developed in Section 3. We first proceed by fitting the directly measured quantities. Using the notation of Section 3.1, the vector $\vec{d}_1$ and its associated covariance matrix $\vec{V}_{d_1}$ are:

$$\vec{d}_1 = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} = \begin{pmatrix} 2.50 \\ 1.67 \\ 1.00 \end{pmatrix}, \ \vec{V}_{d_1} = \begin{pmatrix} 0.0225 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.09 \end{pmatrix}. \tag{5.8}$$

Inverting the identities (5.1) we get:

$$\begin{aligned} A &= \phi(X,C) = X + C \\ B &= \varphi(X,C) = X + \frac{2C}{3} \end{aligned} \tag{5.9}$$

from which we obtain for the matrix $\vec{\tilde{U}}^{(k)}$:

$$\vec{\tilde{U}}^{(k)} = \begin{pmatrix} 1 & 1 \\ 1 & \frac{2}{3} \\ 0 & 1 \end{pmatrix}. \tag{5.10}$$

Because the relations (5.1) are linear our least square problem (3.12) is a linear one, since:

$$\vec{\tilde{D}}^{(k)} = \vec{\tilde{U}}^{(k)} \vec{\tilde{P}}^{(k)} \tag{5.11}$$

and (3.13) yields the least square estimates:

$$\vec{\hat{P}} = \begin{pmatrix} \hat{X} \\ \hat{C} \end{pmatrix} = \begin{pmatrix} 0.8823 \\ 1.3529 \end{pmatrix}, \ \vec{\hat{V}} = \begin{pmatrix} 0.04765 & -0.05293 \\ -0.05293 & 0.06880 \end{pmatrix}. \tag{5.12}$$

From these estimates for X and C we can readily calculate, using (5.9), the least square estimates for A and B:

$$\begin{aligned} \hat{A} &= \hat{X} + \hat{C} = 2.2352 \\ \hat{B} &= \hat{X} + \frac{2\hat{C}}{3} = 1.7842 \end{aligned} \tag{5.13}$$

and by straight forward "propagation of the errors":

$$\begin{aligned} Var(\hat{A}) &= Var(\hat{X}) + Var(\hat{C}) + 2Cov(\hat{X},\hat{C}) \\ Var(\hat{B}) &= Var(\hat{X}) + \frac{4\,Var(\hat{C})}{9} + \frac{4\,Cov(\hat{X},\hat{C})}{3} \\ Cov(\hat{A},\hat{X}) &= Var(\hat{X}) + Cov(\hat{X},\hat{C}) \\ Cov(\hat{B},\hat{C}) &= Var(\hat{X}) + \frac{2\,Cov(\hat{X},\hat{C})}{3} \\ Cov(\hat{A},\hat{C}) &= Var(\hat{C}) + Cov(\hat{X},\hat{C}) \\ Cov(\hat{B},\hat{C}) &= \frac{2\,Var(\hat{C})}{3} + Cov(\hat{X},\hat{C}) \\ Cov(\hat{A},\hat{B}) &= Var(\hat{X}) + \frac{2\,Var(\hat{C})}{3} + \frac{5Cov(\hat{X},\hat{C})}{3} \end{aligned} \tag{5.14}$$

The numerical results are collected in Table 1.

Table 1. RESULTS OF EXAMPLE 5.1

Estimate	Std. Dev.	Correlation Matrix			
$\hat{X}$ =0.8823	0.2183	1.00			
$\hat{A}$ =2.2352	0.1029	-0.23	1.00		
$\hat{B}$ =1.7842	0.0875	0.65	0.59	1.00	
$\hat{C}$ =1.3529	0.2623	-0.92	0.59	-0.31	1.00

The problem will now be solved by fitting the derived quantities. Using the notation of Section 3.2, the vector $\vec{g}_1$ is:

$$\vec{g}_1 = \begin{pmatrix} x_1 \\ x_2 \\ c_1 \end{pmatrix} = \begin{pmatrix} 1.50 \\ 1.00 \\ 1.00 \end{pmatrix}, \tag{5.15}$$

and from (5.1) we obtain for the matrix $\vec{\vec{S}}^{(k)}$:

$$\vec{\vec{S}}^{(k)} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2/3 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.16}$$

Applying (3.25) with the matrices (5.8) and (5.16), the covariance matrix associated with the derived vector $\vec{g}_1$, (5.15), is:

$$\begin{pmatrix} 0.1125 & 0.0600 & -0.0900 \\ 0.0600 & 0.0500 & -0.0600 \\ -0.0900 & -0.0600 & 0.0900 \end{pmatrix} \tag{5.17}$$

Substituting (5.15) and (5.17) into (3.28) yields very precisely the results (5.12) and consequently those of Table 1.

5.2 NONLINEAR FUNCTIONS: $F(A, C) = f(A, C)$.

This example is one that was proposed by R. W. Peelle in the form of a puzzle and has been discussed informally by some evaluators in the nuclear data community. As of this writing, we do not think that there is unanimity as to the root cause of the "puzzle", or if, in fact, this example leads to a "puzzle". The functional relation is:

$$X = F(A,C) = A/C \tag{5.18}$$

and the three independent measurements, with their associated standard deviations, are:

$$\begin{aligned} a_1 &= 1.50 \pm 0.15 \\ a_2 &= 1.00 \pm 0.10 \\ c_1 &= 1.00 \pm 0.20 \end{aligned} \tag{5.19}$$

We first solve this problem fitting the directly measured quantities.

From Section 3.1, the result for $\hat{A}$ is given by (3.16):

$$\hat{A} = a_w = 1.1538 \pm 0.0832 \tag{5.20}$$

and from (3.18)

$$\begin{aligned} \hat{C} &= c_1 = 1.00 \pm 0.20 \\ \hat{X} &= \frac{\hat{A}}{\hat{C}} = 1.1538 \end{aligned} \tag{5.21}$$

The final numerical results are presented in Table 2a.

Table 2a. RESULTS WHEN FITTING DIRECTLY OBSERVED DATA

Estimate	Std. Dev.	Correlation Matrix		
$\hat{X}$ = 1.1538	0.2453	1.00		
$\hat{A}$ = 1.1538	0.0832	0.34	1.00	
$\hat{C}$ = 1.0000	0.2000	-0.94	0	1.00

We now solve this problem by the method we referred to in Section 3.2 as: "from fitting the derived quantities." As we showed in Section 3.2, this method should yield precisely the same results as when the directly measured quantities are fitted. Our main reasons for providing such a numerical example are: (1) to illustrate the iteration procedure that one must go through since the covariance matrix of the derived quantities is only approximate and, (2) to show that a contradiction may arise if one fails to be consistent in the computation of this approximate covariance matrix.

Since in this example $F(A, C)$ is nonlinear, we must use a linear expansion about some estimates $\hat{A}^{(k)}$ and $\hat{C}^{(k)}$, which imply an estimate for X, to calculate the approximate derived data vector $\vec{g}_1^{(k)}$ and its associated covariance matrix $\hat{\vec{V}}_{g_1}^{(k)}$, that enter into the quadratic expression (3.27) which must be minimized and iterated until convergence. For the derived data vector $\vec{g}_1^{(k)}$ we have:

$$\vec{g}_1^{(k)} = \begin{pmatrix} x_1^{(k)} \\ \\ x_2^{(k)} \\ \\ c_1 \end{pmatrix} = \begin{pmatrix} \dfrac{a_1}{\hat{C}^{(k)}} + \hat{X}^{(k)}\left(1 - \dfrac{c_1}{\hat{C}^{(k)}}\right) \\ \dfrac{a_2}{\hat{C}^{(k)}} + \hat{X}^{(k)}\left(1 - \dfrac{c_1}{\hat{C}^{(k)}}\right) \\ c_1 \end{pmatrix} \quad , \tag{5.22}$$

and since $\hat{C}^{(k)}$ will be c_1 for $k \geq 1$, the derived data vector $\vec{g}_1^{(k)}$ will be identical to the derived data vector $\vec{g}_1$. The results of such an iterated procedure, starting form the arbitrary values of 10 and 20 for $\hat{X}^{(0)}$ and $\hat{C}^{(0)}$ respectively are presented into Table 2b.

Table 2b. ITERATION RESULTS FROM FITTING THE DERIVED DATA

			Standard deviations		Corr. Coef.
k	$\hat{X}^{(k)}$	$\hat{C}^{(k)}$	$\hat{X}^{(k)}$	$\hat{C}^{(k)}$	
0	10.0	20.0			
1	9.5577	1.00	0.1001	0.20	-0.9991
2	1.1538	1.00	1.9133	0.20	-0.9991
3	1.1538	1.00	0.2453	0.20	-0.9407
4	1.1538	1.00	0.2453	0.20	-0.9407

These results are in full agreement with those in Table 2a. Based upon these results, the approximate covariance matrix $\hat{\vec{V}}_{g_1}$ that one should associate with the derived data vector $\vec{g}_1$ is:

$$9\hat{\vec{V}}_{g_1} = \begin{pmatrix} 0.07577 & & \\ 0.05377 & 0.06327 & \\ -0.04616 & -0.04616 & 0.04 \end{pmatrix} \tag{5.23}$$

We now explain what was the "puzzle" that R. W. Peelle constructed based upon this example. Let us assume that we are given only the independent measurements a_1 and c_1, with their associated standard deviations. We could construct a derived data vector whose components are x_1 and c_1. What would be the covariance matrix associated with such a derived data vector? Because we would then have single measurements for A and C, their least square estimates would be:

$$\hat{A} = a_1 , \qquad (5.24)$$
$$\hat{C} = c_1 .$$

Consequently:

$$Var\ (x_1) = \frac{Var\ (a_1)}{(\hat{C})^2} + \frac{(\hat{A})^2\ Var(c_1)}{(\hat{C})^4} \qquad (5.25)$$

$$= \frac{Var\ (a_1)}{c_1^2} + \frac{a_1^2\ Var\ (c_1)}{c_1^4} \qquad (5.26)$$

and:

$$Cov(x_1,c_1) = -\frac{\hat{A}\ Var(c_1)}{(\hat{C})^2} \qquad (5.27)$$

$$= -\frac{a_1\ Var(c_1)}{c_1^2} \qquad (5.28)$$

Of course, a covariance matrix calculated with (5.26) and (5.28) is only approximate and based upon the least squares estimates (5.24).

Let us assume now that we are given only a_2 and c_1, rather than a_1 and c_1. We would construct a derived data vector with components x_2 and c_1. Now, instead of(5.24) we would have:

$$\hat{A} = a_2 , \qquad (5.29)$$
$$\hat{C} = c_1 .$$

The covariance matrix associated with the derived data vector whose components are x_2 and c_1 has elements based upon (5.29):

$$Var\ (x_2) = \frac{Var\ (a_2)}{c_1^2} + \frac{a_2^2\ Var(c_1)}{c_1^4} \tag{5.30}$$

and:

$$Cov(x_2,c_1) = -\frac{a_2\ Var(c_1)}{c_1^2}\ . \tag{5.31}$$

Such a covariance matrix is approximate and based upon the least squares estimates (5.29) for A and C.

Let us now consider the situation where we are given a_1 , a_2 and c_1. We could still construct two derived data vectors having components x_1 , c_1 and x_2 , c_1 respectively. If we still associated with these derived data vectors covariance matrices with components given by (5.26) , (5.28) and (5.30), (5.31) respectively these two covariance matrices would be inconsistent with each other. This is because these two covariance matrices are based upon different least squares estimates for A, unless a_1 and a_2 are numerically the same. Such an inconsistency is what gave rise to Peelle's puzzle. Quite specifically, what is done to obtain Peelle's puzzle is that one associates with the derived data vector $\vec{g}_1$ a covariance matrix with elements given by:

$$
\begin{aligned}
Var(x_i) &= \frac{Var(a_i)}{c_1^2} + \frac{a_i^2\, Var(c_1)}{c_1^4} \\
Cov(x_1, x_2) &= a_1 a_2\, Var(c_1) \\
Cov(x_i, c_1) &= -\frac{a_i\, Var(c_1)}{c_1^2}
\end{aligned} \tag{5.32}
$$

Furthermore, one does not consider that such a covariance matrix is approximate, by being conditional upon some least square estimate for A because X is a nonlinear function of A and C. Substituting numerical values (5.19) into (5.32) one obtains the covariance matrix:

$$
\vec{V}_{g_1} = \begin{pmatrix} 0.1125 & & \\ 0.06 & 0.05 & \\ -0.06 & -0.04 & 0.04 \end{pmatrix} \tag{5.33}
$$

This covariance matrix is very different from what we claim is the correct approximate one given by (5.23). If one associates with the vector $\vec{g}_1$, the covariance matrix (5.33) and one seeks least squares estimates for X and C by minimizing the quadratic expression (3.19), then one solves a linear least squares problem whose results are given in Table 2c.

Table 2c. RESULTS FROM PEELLE'S PUZZLE.

Estimate	Std. Dev.	Correlation Matrix		
$\hat{X}$ = 0.882	0.213	1.00		
$\hat{A}$ = 1.090	0.141	0.34	1.00	
$\hat{C}$ = 1.235	0.175	-0.92	-0.66	1.00

These results are different from those in Table 2a and were said by R. W. Peelle to give rise to a puzzle. The puzzle was: Why do we not get the same results when we fit the derived data and the directly observed data? Peelle thought that the results given in Table 2a were correct, and that some mistake must have been made in obtaining the results in Table 2c.

5.3 NONLINEAR FUNCTIONS: $F(A,C) \neq f(B,C)$.

We now consider the functional relations:

$$\begin{aligned} X &= F(A,C) = A - C \\ X &= f(B,C) = \frac{B}{C} \end{aligned} \tag{5.34}$$

and their inverse:

$$\begin{aligned} A &= \phi(X,C) = X + C \\ B &= \varphi(X,C) = X\,C \end{aligned} \tag{5.35}$$

with the three independent measured data:

$$\begin{aligned} a_1 &= 2.50 \pm 0.05 \\ b_1 &= 1.00 \pm 0.30 \\ c_1 &= 1.00 \pm 0.30 \end{aligned} \tag{5.36}$$

We can obtain the least squares estimates that we seek by solving the traditional nonlinear least squares problem (3.12), what we refer to as from fitting the directly observed data. The converged results of such a procedure are given in Table 3a.

Table 3a. RESULTS WHEN FITTING THE DIRECTLY OBSERVED DATA

Estimate	Std. Dev.	Correlation Matrix			
$\hat{X}$ = 1.783	0.216	1.00			
$\hat{C}$ = 0.712	0.206	-0.97	1.00		
$\hat{A}$ = 2.495	0.050	0.31	-0.09	1.00	
$\hat{B}$ = 1.269	0.220	-0.92	0.99	0.08	1.00

Alternately we can proceed by fitting the derived data. Because of the nonlinearity of the functional dependence of X upon B and C, we must linearize it via an expansion about some estimates $\hat{B}^{(k)}$ and $\hat{C}^{(k)}$ and iterate the solution. That is to say, we minimize the quadratic expression (3.27) iterating until convergence is achieved. The expression for $\vec{g}_1^{(k)}$ is:

$$\vec{g}_1^{(k)} = \begin{pmatrix} x_1 \\ x_2^{(k)} \\ c_1 \end{pmatrix} = \begin{pmatrix} a_1 - c_1 \\ \dfrac{\hat{B}^{(k)}}{\hat{C}^{(k)}} + \dfrac{(b_1 - \hat{B}^{(k)})}{\hat{C}^{(k)}} - (c_1 - \hat{C}^{(k)}) \dfrac{\hat{B}^{(k)}}{\hat{C}^{(k)2}} \\ c_1 \end{pmatrix} \tag{5.37}$$

and for $\vec{\vec{V}}_{g_1}^{(k)}$ it is:

$$\vec{V}_{g_1}^{(k)} = \begin{pmatrix} Var(a_1) + Var(c_1) & & \\ \dfrac{\hat{B}^{(k)}\, Var(c_1)}{\hat{C}^{(k)^2}} & \dfrac{Var(b_1)}{\hat{C}^{(k)}} + \hat{B}^{(k)^2}\, \dfrac{Var(c_1)}{\hat{C}^{(k)^4}} & \\ -Var(c_1) & -\dfrac{\hat{B}^{(k)}\, Var(c_1)}{\hat{C}^{(k)^2}} & Var(c_1) \end{pmatrix} \tag{5.38}$$

The results of such a procedure are identical to those given in Table 3a. It is important to note that in this problem if instead of (5.37) we take the derived data vector to be:

$$\vec{g}_1 = \begin{pmatrix} a_1 - c_1 \\ b_1 / c_1 \\ c_1 \end{pmatrix} \quad , \tag{5.39}$$

that is to say we do not use the same linearization to calculate both $\vec{g}_1$ and $\vec{V}_{g_1}^{(k)}$ in the least square procedure, then we do not obtain the results given in Table 3a. We must, therefore, emphasize that consistency between the least squares results obtained by fitting the directly observed data and the derived data is only guaranteed if the same linearization of the nonlinear functional dependences are used to calculate the elements of the derived data vector and its associated covariance matrix.

5.4 CURVE FITTING.

Finally we consider an example of performing a least squares fit to derived quantities. We have the directly measured data are given in Table 4.

Table 4. DIRECTLY MEASURED DATA.

i	1	2	3	4	5	6	7	8
E_i	0.8	1.0	2.3	3.4	4.5	7.4	8.8	9.7
$a_1(i)$	19	30	27	41	52	53	63	78

We assume that all of the data $a_1(i)$ in Table 4 are independent and have a 10% standard deviation associated with them. We also assume that the uncertainties in the independent variable denoted E_i can be neglected. Furthermore, we have the independent measurement of the quantity C:

$$c_1 = 1.0 \pm 0.2 \tag{5.40}$$

We are interested in the derived quantity:

$$X(E) = F\,(A(E),\, C) = A(E) \star C \tag{5.41}$$

Therefore, we have the derived data:

$$x_1(i) = a_1(i) * c_1 \tag{5.42}$$

We assume that we also have:

$$X(E) = R(E; \vec{H}\,) = H_1 + H_2 E \tag{5.43}$$

and that we seek a least squares estimate for the parameters H_1 and H_2. The traditional method of solving this problem would be by fitting the directly observed data. That is to say, minimize the quadratic expression corresponding to (4.23) where we have:

$$\vec{d}_1 - \vec{D} = \begin{pmatrix} a_1(1) - \dfrac{(H_1 + H_2E_1)}{C} \\ a_1(2) - \dfrac{(H_1 + H_2E_2)}{C} \\ \vdots \\ a_1(8) - \dfrac{(H_1 + H_2E_8)}{C} \\ c_1 - C \end{pmatrix} \tag{5.44}$$

and:

$$\vec{V}_{d_1} = \begin{pmatrix} Var(a_1(1)) & 0 & \cdots & 0 & 0 \\ 0 & Var(a_1(2)) & \cdots & 0 & 0 \\ 0 & 0 & \vdots & \vdots & \vdots \\ 0 & 0 & \cdot & Var(a_1(8)) & 0 \\ 0 & 0 & \cdots & 0 & Var(c_1) \end{pmatrix} \tag{5.45}$$

This nonlinear least squares problem is solved by iteration and yields the results in Table 4a and by the curve labelled A in figure 1.

Table 4a. RESULTS WHEN FITTING DIRECTLY OBSERVED DATA.

Estimate	Std. Dev.	Correlation Matrix		
$\hat{H}_1$ = 17.12	3.819	1.00		
$\hat{H}_2$ = 5.689	1.244	0.69	1.00	
$\hat{C}$ = 1.000	0.20	0.90	0.91	1.00

This problem can also be solved by fitting the derived data, i.e. by minimizing with respect to H_1, H_2 and C the quadratic expression (4.13). However, we must emphasize again that the results given in Table 4a will only be obtained if a consistent linearization is used to calculate the vector we denote $\vec{g}_1^{(k)}$ and the elements of the covariance matrix $\vec{V}_{g_1}^{(k)}$. Given informal discussions that have taken place concerning ***PPP***, we conjecture that a number of evaluators might attempt to fit the derived data in the following manner. The following vector would be taken to be the derived data to be fitted:

$$\vec{g}_1 = \begin{pmatrix} a_1(1) \;\; c_1 \\ a_1(2) \;\; c_1 \\ \vdots \\ a_1(8) \;\; c_1 \\ c_1 \end{pmatrix} , \tag{5.46}$$

and the elements of its associated covariance matrix taken to be given by:

$$\begin{aligned} \vec{V}_{g_1}(i, i) &= c_1^2 \; Var\,(a_1\,(i)) + a_i^2 \; Var\,(c_1) &&, \; for\ i \le 8 \\ \vec{V}_{g_1}\,(i,j) &= a_i \, a_j \; Var\,(c_1) &&, \; for\ i\ and\ j \le 8 \\ \vec{V}_{g_1}\,(9, 9) &= Var\,(c_1) &&, \; and \\ \vec{V}_{g_1}\,(i, 9) &= a_i \; Var\,(c_1) &&, \; for\ i \le 8 \quad . \end{aligned} \tag{5.47}$$

The derived data vector (5.46) and its associated covariance matrix (5.47) would be treated as if they were directly measured data and the following quadratic expression would be minimized with respect to the components of the vector $\boldsymbol{P}$:

$$(\vec{g}_1 - \vec{T}\vec{P})^t \, \vec{V}_{g_1}^{-1} \, (\vec{g}_1 - \vec{T}\vec{P}) \tag{5.48}$$

where:

$$\vec{T} \equiv \begin{pmatrix} 1 & E_1 & O \\ 1 & E_2 & 0 \\ \vdots & \vdots & \vdots \\ 1 & E_8 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.49}$$

and

$$\vec{P} \equiv \begin{pmatrix} H_1 \\ H_2 \\ C \end{pmatrix} \tag{5.50}$$

If one were indeed to solve this problem in that manner the results would be as given in Table 4b and the curve labelled B in Figure 1, rather than those given in Table 4a and the curve labelled A.

Table 4b. RESULTS WHEN TREATING DERIVED DATA AS IF THEY WERE DIRECTLY OBSERVED DATA.

Estimate	Std. Dev.	Correlation Matrix		
$\hat{H}_1$ = 10.47	3.167	1.00		
$\hat{H}_2$ = 3.478	1.002	0.55	1.00	
$\hat{C}$ = 0.611	0.156	0.85	0.87	1.00

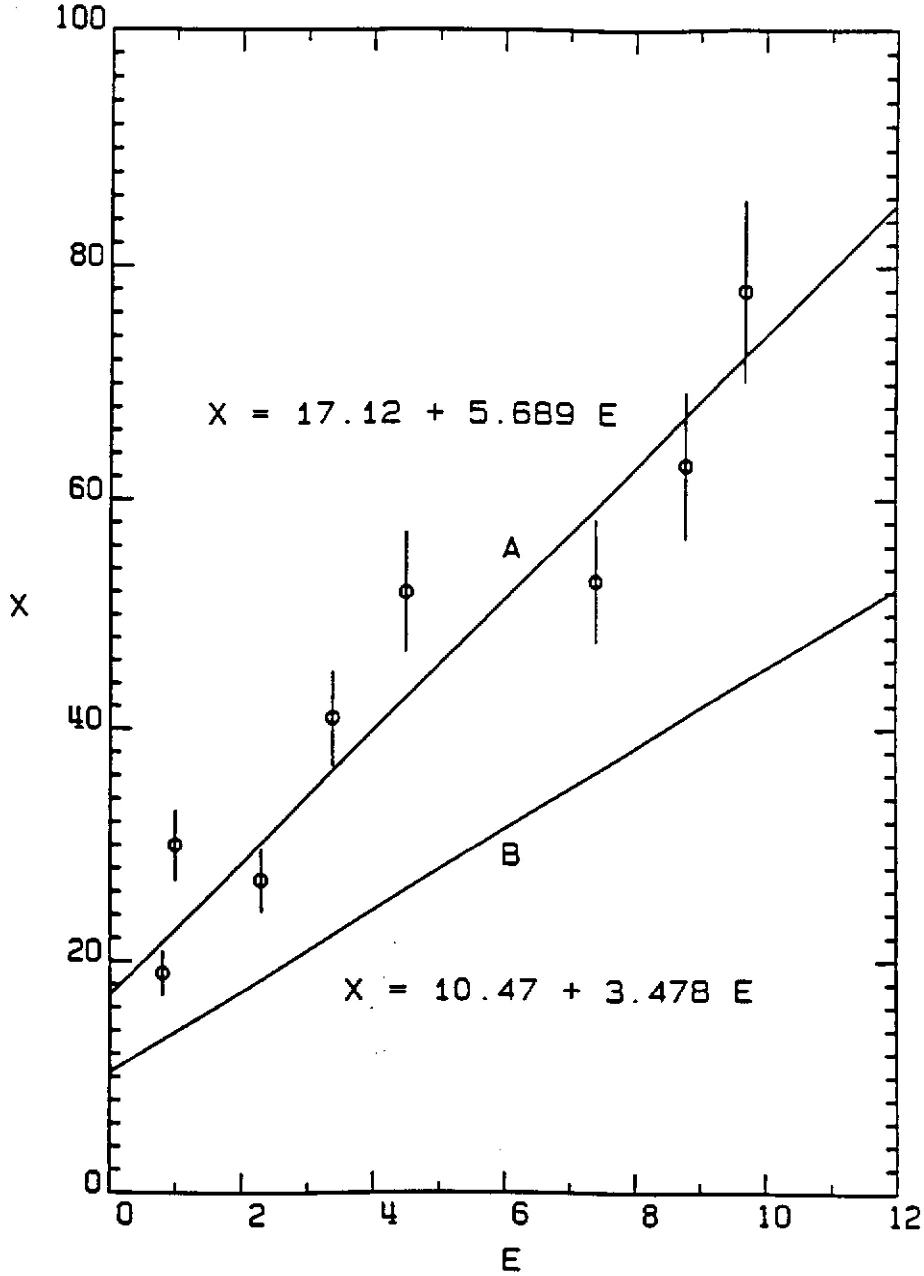

Figure 1. Example 5.4: Curve fitting.

6.0 CONCLUSION

In this report we have analyzed the general problem of generating the covariance matrix associated with quantities derived from directly measured data via nonlinear relations, and the use of such matrices in data evaluations using the least squares method. When the derived quantities are nonlinear functions of the measured data, which is very often the case in practice, we can only obtain an approximate covariance matrix associated with the derived quantities. These approximate covariance matrices depend implicitly upon estimated values for the measured data. We have analyzed a few situations, based upon a problem suggested by R. W. Peelle, where inconsistencies can occur in evaluation work if one ignores the fact that these covariance matrices are not only approximate but a function of the very estimates one is seeking in the evaluation work. We have indicated how these inconsistencies can be avoided by a suitable iteration procedure.

REFERENCE

1. R. W. Peelle. "Peelle's Pertinent Puzzle," private communication (October 1987) (Revised October 28, 1988). See also D. L. Smith, "Probability, Statistics, and Data Uncertainties in Nuclear Science and Technology," (*American Nuclear Society,* La Grange Park, IL, 1991) p. 205ff.

INTERNAL DISTRIBUTION

1. B. R. Appleton
2. J. K. Dickens
3. C. Y. Fu
4. T. A. Gabriel
5. J. A. Harvey
6. D. M. Hetrick

7-8. D. C. Larson

9. N. M. Larson
10. R. W. Peelle

11-15. F. G. Perey

16. R. W. Roussin
17. R. C. Ward
18. L. W. Weston
19. R. W. Brockett (consultant)
20. J. J. Dorning (consultant)
21. J. E. Leiss (consultant)
22. N. Moray (consultant)
23. M. F. Wheeler (consultant)
24. EPMD Reports Office

25-27. Laboratory Records Department

28. Laboratory Records, ORNL-RC
29. Document Reference Section
30. Central Research Library
31. ORNL Patent Section

EXTERNAL DISTRIBUTION

32. Office of the Assistant Manager for Energy Research and Development, U. S. Department of Energy, Oak Ridge Operations, P. O. Box 2008, Oak Ridge, TN 37831.
33. S. L. Whetstone, Division of Nuclear Physics, ER 23/GTN, U. S. Department of Energy, Washington, D. C. 20585.
34. R. A. Meyer, Division of Nuclear Physics, ER 23/GTN, U. S. Department of Energy, Washington, D. C. 20585.
35. S. E. Berk, Office of Energy Research, Reactors and Systems Radiation, U. S. Department of Energy, Washington, D. C. 20585.
36. Jincheng Xu, Director, Department of Physics, Institute of Atomic Energy, P. O. Box 275, Beijing, P. R. China.

37-41. Zhi-xiang Zhao, Chinese Nuclear Data Center, P. O. Box 275(41), Beijing, P. R. China.

42-43. National Nuclear Data Center, ENDF, Brookhaven National Laboratory, Upton, N.Y. 11973.

44. Chief, Mathematics and Geoscience Branch, U. S. Department of Energy, Washington, D. C. 20585.

45-55. Office of Scientific and Technical Information, P. O. Box 62, Oak Ridge, TN 37831.

56. C. L. Dunford, NNDC, Building 197-D, Brookhaven National Laboratory, Upton, NY 11973.
57. W. P. Poenitz, Argonne National Laboratory-West, P. O. Box 2528, Idaho Falls, ID 83403.
58. P. G. Young, Theory Division, T-2, MS-B243, Los Alamos National Laboratory, Los Alamos, NM 87543.
59. D. L. Smith, Building 314, Argonne National Laboratory, Argonne, IL 60439.
60. F. M. Mann, L7-29, Hanford Engineering Development Laboratory, P. O. Box 1970, Richland, WA 99352.
61. R. E. Schenter, L7-29, Hanford Engineering Development Laboratory, P. O. Box 1970, Richland, WA 99352.
62. S. Chiba, Japan Nuclear Data Center, JAERI, Tokai Research Establishment, Tokai-Mura, Naka-Gun, IBARAKI-KEN, Japan.
63. F. H. Froehner, Institute fur Neutronenphysik und Reaktortechnik, Kernforschungszentrum, Karlsruhe, Postfach 3640, D-7500, Karlsruhe, Germany.

64. H. Gruppelaar, Department of Physics, Netherlands Energy Research Foundation (ECN), P. O. Box 1, 3 Westerduinweg, NL-1755 ZG Petten, Netherlands.
64. H. K. Vonach, Institute Fuer Radiumforschung und Kernphysik, Boltzmanngasse 3, A-1090, Wien, Austria.
65. Douglas W. Muir, Los Alamos National Laboratory, T2, MS-B243, Los Alamos, New Mexico 87545.
66. J. J. Schmidt, International Atomic Energy Agency/Nuclear Data Section, P. O. Box 100, A-1400, Vienna, Austria.
67. B. H. Patrick, Building 418, U. K. Atomic Energy Authority, Atomic Energy Research Establishment, Harwell, Didcot, Oxon, OX11 ORA, United Kingdom.
68. Y. Kanda, Department of Energy Conversion Engineering, Kyushu University, Kasuga, Fukuoka 816, Japan.
69. Delin Zhou, P. O. Box 275(41), Beijing, P. R. China.
70. Zuxun Sun, President, Institute of Atomic Energy, P. O. Box 275, Beijing, P. R. China.
71. Claes Nordborg, OEDC/NEA Data Bank, Batiment 445, F-91191 Gif/Yvette, Cedex, France.
72. Y. Kikuchi, Nuclear Data Center, Department of Physics, JAERI, Tokai-mura, Naka-gun, Ibaraki-ken 319-11, Japan.
73. V. Pronyaev, Institute of Physics and Power Engineering, 249020 Obininsk, Kaluga Region, U.S.S.R.
74. Claude Bastian, CEC-IRC, Central Bureau for Nuclear Measurements, B-2440 Geel, Belgium.
75. Wolf Mannhart, Physikalisch-Technische Bundesanstalt, Lab. 7.23, Bundesallee 100 W-3300 Braunschweig, Germany.
76. Frank Schmittroth, Hanford Engineering Development Laboratory, Westinghouse Hanford Company, L7-29, P. O. Box 1970, Richland, Washington 99352.
77. Roger M. White, Physics Department, MS L-298, Lawrence Livermore National Laboratory, P. O. Box 808, Livermore, CA 94550.
78. Jehudah Wagschal, Racah Institute of Physics, Hebrew University, Jerusalem, Israel.
79. Y. Yeivin, Racah Institute of Physics, The Hebrew University, IL-91904, Jerusalem, Israel.
80. D. R. Weaver, School of Physics and Space Research, University of Birmingham, P. O. Box 363, Birmingham B15 2TT, United Kingdom.
81. Michael C. Moxon, 3 Hyde Copse, Marcham, Abingdon OX136TT, Oxon, England.