ANL/NDM--62

DE82 007201

ANL/NDM-62

COVARIANCE MATRICES AND APPLICATIONS TO THE FIELD OF NUCLEAR DATA*

by

Donald L. Smith

MASTER

November 1981

DISCLAIMER

*This work supported by the U.S. Department of Energy.

Applied Physics Division
Argonne National Laboratory
9700 South Cass Avenue
Argonne, Illinois 60439
USA

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

NUCLEAR DATA AND MEASUREMENTS SERIES

The Nuclear Data and Measurements Series presents results of studies in the field of microscopic nuclear data. The primary objective is the dissemination of information in the comprehensive form required for nuclear technology applications. This Series is devoted to: a) measured microscopic nuclear parameters, b) experimental techniques and facilities employed in measurements, c) the analysis, correlation and interpretation of nuclear data, and d) the evaluation of nuclear data. Contributions to this Series are reviewed to assure technical competence and, unless otherwise stated, the contents can be formally referenced. This Series does not supplant formal journal publication but it does provide the more extensive information required for technological applications (e.g., tabulated numerical data) in a timely manner.

OTHER ISSUES IN THE ANL/NDM SERIES ARE:

ANL/NDM-1 Cobalt Fast Neutron Cross Sections-Measurement and Evaluation by P. T. Guenther, P. A. Moldauer, A. B. Smith, D. L. Smith and J. F. Whalen, July 1973.

ANL/NDM-2 Prompt Air-Scattering Corrections for a Fast-Neutron Fission Detector: $E_n \leq 5$ MeV by Donald L. Smith, September 1973.

ANL/NDM-3 Neutron Scattering from Titanium; Compound and Direct Effects by E. Barnard, J. deVilliers, P. Moldauer, D. Reitmann, A. Smith and J. Whalen, October 1973.

ANL/NDM-4 ^{90}Zr and ^{92}Zr; Neutron Total and Scattering Cross Sections by P. Guenther, A. Smith and J. Whalen, July 1974.

ANL/NDM-5 Delayed Neutron Data - Review and Evaluation by Samson A. Cox, April 1974.

ANL/NDM-6 Evaluated Neutronic Cross Section File for Niobium by R. Howerton, Lawrence Livermore Laboratory and A. Smith, P. Guenther and J. Whalen, Argonne National Laboratory, May 1974.

ANL/NDM-7 Neutron Total and Scattering Cross Sections of Some Even Isotopes of Molybdenum and the Optical Model by A. B. Smith, P. T. Guenther and J. F. Whalen, June 1974.

ANL/NDM-8 Fast Neutron Capture and Activation Cross Sections of Niobium Isotopes by W. P. Poenitz, May 1974.

ANL/NDM-9 Method of Neutron Activation Cross Section Measurement for E_n = 5.5-10 MeV Using the D(d,n)He-3 Reaction as a Neutron Source by D. L. Smith and J. W. Meadows, August 1974.

ANL/NDM-10 Cross Sections for (n,p) Reactions on ^{27}Aℓ, ^{46,47,48}Ti, ^{54,56}Fe, ^{58}Ni, ^{59}Co and ^{64}Zn from Near Threshold to 10 MeV by Donald L. Smith and James W. Meadows, January 1975.

ANL/NDM-11 Measured and Evaluated Fast Neutron Cross Sections of Elemental Nickel by P. Guenther, A. Smith, D. Smith and J. Whalen, Argonne National Laboratory and R. Howerton, Lawrence Livermore Laboratory, July 1975.

ANL/NDM-12 A Spectrometer for the Investigation of Gamma Radiation Produced by Neutron-Induced Reactions by Donald L. Smith, April 1975.

ANL/NDM-13 Response of Several Threshold Reactions in Reference Fission Neutron Fields by Donald L. Smith and James W. Meadows, June 1975.

ANL/NDM-14 Cross Sections for the ^{66}Zn(n,p)^{66}Cu, ^{113}In(n,n')^{113m}In and ^{115}In(n,n')^{115m}In Reactions from Near Threshold to 10 MeV by Donald L. Smith and James W. Meadows, July 1975.

ANL/NDM-15 Radiative Capture of Fast Neutrons in ^{165}Ho and ^{181}Ta by W. P. Poenitz, June 1975.

ANL/NDM-16 Fast Neutron Excitation of the Ground-State Rotational Band of ^{238}U by P. Guenther, D. Havel and A. Smith, September 1975.

ANL/NDM-17 Sample-Size Effects in Fast-Neutron Gamma-Ray Production Measurements: Solid-Cylinder Samples by Donald L. Smith, September 1975.

ANL/NDM-18 The Delayed Neutron Yield of ^{238}U and ^{241}Pu by J. W. Meadows January 1976.

ANL/NDM-19 A Remark on the Prompt-Fission-Neutron Spectrum of ^{252}Cf by P. Guenther, D. Havel, R. Sjoblom and A. Smith, March 1976.

ANL/NDM-20 Fast-Neutron-Gamma-Ray Production from Elemental Iron: $E_n \lesssim$ 2 MeV by Donald L. Smith, May 1976.

ANL/NDM-21 Note on the Experimental Determination of the Relative Fast-Neutron Sensitivity of a Hydrogenous Scintillator by A. Smith, P. Guenther and R. Sjoblom, June 1976.

ANL/NDM-22 Note on Neutron Scattering and the Optical Model Near A=208 by P. Guenther, D. Havel and A. Smith, September 1976.

ANL/NDM-23 Remarks Concerning the Accurate Measurement of Differential Cross Sections for Threshold Reactions Used in Fast-Neutron Dosimetry for Fission Reactors by Donald L. Smith, December 1976.

ANL/NDM-24 Fast Neutron Cross Sections of Vanadium and an Evaluated Neutronic File by P. Guenther, D. Havel, R. Howerton, F. Mann, D. Smith, A. Smith and J. Whalen, May 1977.

ANL/NDM-25 Determination of the Energy Scale for Neutron Cross Section Measurements Employing a Monoenergetic Accelerator by J. W. Meadows, January 1977.

ANL/NDM-26 Evaluation of the IN-115(N,N')IN-115M Reaction for the ENDF/B-V Dosimetry File by Donald L. Smith, December 1976.

ANL/NDM-27 Evaluated (n,p) Cross Sections of ^{46}Ti, ^{47}Ti and ^{48}Ti by C. Philis and O. Bersillon, Bruyeres-le-Chatel, France and D. Smith and A. Smith, Argonne National Laboratory, January 1977.

ANL/NDM-28 Titanium-II: An Evaluated Nuclear Data File by C. Philis, Centre d'Etudes de Bruyéres-le-Châtel, R. Howerton, Lawrence Livermore Laboratory and A. B. Smith, Argonne National Laboratory, June 1977.

ANL/NDM-29 Note on the 250 keV Resonance in the Total Neutron Cross Section of ^{6}Li by A. B. Smith, P. Guenther, D. Havel and J. F. Whalen, June 1977.

ANL/NDM-30 Analysis of the Sensitivity of Spectrum-Average Cross Sections to Individual Characteristics of Differential Excitation Functions by Donald L. Smith, March 1977.

ANL/NDM-31 Titanium-I: Fast Neutron Cross Section Measurements by P. Guenther, D. Havel, A. Smith and J. Whalen, May 1977.

ANL/NDM-32 Evaluated Fast Neutron Cross Section of Uranium-238 by W. Poenitz, E. Pennington, and A. B. Smith, Argonne National Laboratory and R. Howerton, Lawrence Livermore Laboratory, October 1977.

ANL/NDM-33 Comments on the Energy-Averaged Total Neutron Cross Sections of Structural Materials by A. B. Smith and J. F. Whalen, June 1977.

ANL/NDM-34 Graphical Representation of Neutron Differential Cross Section Data for Reactor Dosimetry Applications by Donald L. Smith, June 1977.

ANL/NDM-35 Evaluated Nuclear Data File of Th-232 by J. Meadows, W. Poenitz, A. Smith, D. Smith and J. Whalen, Argonne National Laboratory and R. Howerton, Lawrence Livermore Laboratory, February 1978.

ANL/NDM-36 Absolute Measurements of the ^{233}U(n,f) Cross Section Between 0.13 and 8.0 MeV by W. P. Poenitz, April 1978.

ANL/NDM-37 Neutron Inelastic Scattering Studies for Lead-204 by D. L. Smith and J. W. Meadows, December 1977.

ANL/NDM-38 The Alpha and Spontaneous Fission Half-Lives of ^{242}Pu by J. W. Meadows, December 1977.

ANL/NDM-39 The Fission Cross Section of ^{239}Pu Relative to ^{235}U from 0.1 to 10 MeV by J. W. Meadows, March 1978.

ANL/NDM-40 Statistical Theory of Neutron Nuclear Reactions by P. A. Moldauer, February 1978.

ANL/NDM-41 Energy-Averaged Neutron Cross Sections of Fast-Reactor Structural Materials by A. Smith, R. McKnight and D. Smith, February 1978.

ANL/NDM-42 Fast Neutron Radiative Capture Cross Section of ^{232}Th by W. P. Poenitz and D. L. Smith, March 1978.

ANL/NDM-43 Neutron Scattering from ^{12}C in the Few-MeV Region by A. Smith, R. Holt and J. Whalen, September 1978.

ANL/NDM-44 The Interaction of Fast Neutrons with ^{60}Ni by A. Smith, P. Guenther, D. Smith and J. Whalen, January 1979.

ANL/NDM-45 Evaluation of ^{235}U(n,f) between 100 KeV and 20 MeV by W. P. Poenitz, July 1979.

ANL/NDM-46 Fast-Neutron Total and Scattering Cross Sections of ^{107}Ag in the MeV Region by A. Smith, P. Guenther, G. Winkler and J. Whalen, January 1979.

ANL/NDM-47 Scattering of MeV Neutrons from Elemental Iron by A. Smith and P. Guenther, March 1979.

ANL/NDM-48 ^{235}U Fission Mass and Counting Comparison and Standardization by W. P. Poenitz, J. W. Meadows and R. J. Armani, May 1979.

ANL/NDM-49 Some Comments on Resolution and the Analysis and Interpretation of Experimental Results from Differential Neutron Measurements by Donald L. Smith, November 1979.

ANL/NDM-50 Prompt-Fission-Neutron Spectra of ^{233}U, ^{235}U, ^{239}Pu and ^{240}Pu Relative to that of ^{252}Cf by A. Smith, P. Guenther, G. Winkler and R. McKnight, September 1979.

ANL/NDM-51 Measured and Evaluated Neutron Cross Sections of Elemental Bismuth by A. Smith, P. Guenther, D. Smith and J. Whalen, April 1980.

ANL/NDM-52 Neutron Total and Scattering Cross Sections of ^{6}Li in the Few MeV Region by P. Guenther, A. Smith and J. Whalen, February 1980.

ANL/NDM-53 Neutron Source Investigations in Support of the Cross Section at the Argonne Fast-Neutron Generator by James W. Meadows and Donald L. Smith, May 1980.

ANL/NDM-54 The Nonelastic-Scattering Cross Sections of Elemental Nickel by A. B. Smith, P. T. Guenther and J. F. Whalen, June 1980.

ANL/NDM-55 Thermal Neutron Calibration of a Tritium Extraction Facility using the $^{6}Li(n,t)^{4}He/^{197}Au(n,\gamma)^{198}Au$ Cross Section Ratio for Standardization by M. M. Bretscher and D. L. Smith, August 1980.

ANL/NDM-56 Fast-neutron Interactions with ^{182}W, ^{184}W and ^{186}W by P. T. Guenther, A. B. Smith and J. F. Whalen, June 1981.

ANL/NDM-57 The Total, Elastic- and Inelastic-Scattering of Fast-Neutron Cross Sections of Natural Chromium by Peter T. Guenther, Alan B. Smith and James F. Whalen, January 1981.

ANL/NDM-58 Review of Measurement Techniques for the Neutron Capture Process, Wolfgang P. Poenitz, (to be published).

ANL/NDM-59 Review of the Importance of the Neutron Capture Process in Fission Reactors, Wolfgang P. Poenitz, (to be published).

ANL/NDM-60 Neutron Capture Activation Cross Sections of $^{94,96}Zr$, $^{98,100}Mo$, and $^{110,114,116}Cd$ at Thermal and 30 keV Energy, John M. Wyrick and Wolfgang P. Poenitz, (to be published).

ANL/NDM-61 Fast-neutron Total and Scattering Cross Sections of ^{58}Ni by Carl Budtz-Jørgensen, Peter T. Guenther, Alan B. Smith and James F. Whalen, (to be published).

TABLE OF CONTENTS

		Page
PREFACE		vii
ABSTRACT		ix
I.	INTRODUCTION	1
II.	CURVILINEAR COORDINATES AND METRIC COEFFICIENTS - ORIGIN OF THE FORMALISM	3
III.	ERROR ANALYSIS FOR A SCALAR QUANTITY	9
IV.	ERROR ANALYSIS FOR A VECTOR QUANTITY	15
V.	COMBINATION OF RATIO AND STANDARD ERRORS	22
VI.	SOME COMMENTS RELEVANT TO THE GENERATION OF - AND ERROR PROPAGATION WITH - COVARIANCE MATRICES	25
VII.	COVARIANCE MATRICES AND THE LEAST-SQUARES METHOD FOR SOLVING OVER-DETERMINED SYSTEMS OF LINEAR EQUATIONS	27
VIII.	ONE METHOD FOR GENERATING EVALUATED CROSS SECTIONS FROM A DATA BASE USING THE LEAST-SQUARES METHOD	37
IX.	USE OF AN APRIORI AND LINEARIZATION OF NON-LINEAR LEAST-SQUARES PROBLEMS	41
X.	HOW TO UPDATE AN EVALUATION WITH NEW DATA WITHOUT PERFORMING A COMPLETELY NEW EVALUATION	43
XI.	USE OF THE LEAST-SQUARES METHOD IN DATA FITTING	46
XII.	CONCLUSIONS	62
ACKNOWLEDGMENTS		63
REFERENCES		64

PREFACE

Physicists like to think of themselves as totally rational people, but apparently they are just as prone to fads and fashion in their work as are other mortals. The topic of covariance matrices has emerged as a lively contemporary issue in the field of nuclear data. Furthermore, it is an issue which has evoked strong and seemingly unreconcilable partisan sentiments within this community of researchers.

No one denies that there is a need for improvements in standards for reporting experimental data and their uncertainties, and in the methods used for evaluating nuclear data for applications. The controversy centers mainly around the matter of whether subjective interpretation and judgment should continue to play an important role in data analyses, or whether they should be entirely replaced by scientifically rigorous procedures. Proponents of the scientifically rigorous approach claim, with justification, that subjective methods introduce biases which are nearly impossible to trace. Furthermore, they believe that science is not an art and that scientists should adhere to rigorous methods, without resorting to expediency, otherwise their activities will cease to be science. Those who believe that educated intuition and judgment can never be entirely replaced argue convincingly that the requisite input for rigorous analyses is often not available, or is dubious in quality. They claim that there has always been a place, or even a requirement, for the exercise of good judgment in science. Superimposed upon this controversy is the fact that strict adherence to idealistic standards for reporting data and uncertainties would involve the expenditure of more time, and the management of significantly greater quantities of information, than the levels which are already straining our existing data manipulation institutions. The community must decide whether the need for detailed uncertainty information is sufficiently great to justify the additional effort and costs involved, and then chart an appropriate course for future endeavors.

This report has <u>not</u> been prepared primarily to support a particular partisan position in this issue. During a sabbatical leave (September 1980 to August 1981) at the Central Bureau for Nuclear Measurements, Geel, Belgium, I was challenged by Horst Liskien to develop a rigorous procedure for deriving uncertainties in calculated spectrum-averaged activation cross sections. I glanced through the recent literature and found some papers which indicated the appropriate direction to follow, although they did not impress me as being entirely satisfactory tutorial documents for someone who wants to learn a working knowledge of covariance error analysis and evaluation techniques with minimal expenditure of time. I knew that there would be many books and older documents available on this subject, but I did not have time to seek, sort through and study all of this material. Pragmatic considerations led me to be concerned more with the mechanics of error analysis and less with the mathematical foundations of this field. Essentially, I acquired the information I needed to know to get started from a few papers and then worked out my own procedures during the course of solving specific problems at hand. The notes which I prepared for my own benefit were ultimately collected into a handout distributed preceding a lecture on covariances which I presented at Geel in June 1981. Sections II through X of this report are extracted from these lecture notes. Section XI is based on later work.

No one who has written lately about the topic of covariances in nuclear data applications can make much claim to originality. Although the concepts and formulas are known to have appeared in many places and forms over a time span of nearly two centuries, I believe that there is a need to restate some of these old concepts in the framework of specific contemporary nuclear data applications. This reports reflects one such endeavor.

It is possible to approach this subject in many ways. My objective has been to produce a self-teaching guide to the topic so that a serious reader could learn the mechanics of covariance analysis by studying this report. I have avoided discussing statistical matters to any great extent, so the student will have to seek elucidation on this matter elsewhere. Normal (Gaussian) statistical laws are commonly assumed in Physics applications. Normal distribution laws apply to large samplings and random variables. There are no doubt many situations where the assumption of Normal statistics is a poor approximation. The validity of most of the error propagation formulas is not dependent upon the choice of statistics, but one must insure that the errors assumed for the parameters within a given problem all correspond to the same confidence level. Some understanding of the statistical nature of the parameters is thus required implicitly if one hopes to properly interpret the errors derived from propagation calculations.

What I will present here is my understanding of covariance analysis. I have selected my own notation and examples. I would like to have given even more examples than you will find in this report, but realistic examples which do not lead to extensive numerical work are not easy to conjure. Therefore, in some instances, I have illustrated the principles with symbolic examples; i.e., the required procedures are indicated, but numerical values are not inserted. The level of mathematical sophistication in this treatment is at the undergraduate level. An understanding of coordinate systems, functions, differential calculus and the basic concepts of matrix algebra are all that is needed. I will try to indicate the steps involved fairly completely and will not leave too many details for the reader to fill in.

Hopefully, the content of this report, and other recent contributions, will help nuclear data researchers to address the problem of improving the accuracy of their endeavors in ways which are both practical and scientifically acceptable.

I. INTRODUCTION

The analysis of experimental nuclear data and the evaluation of the data base for applications are matters which have been receiving attention lately. The large expense involved in performing new experiments, the need for researchers to operate with restricted funding, and increasingly stringent data accuracy requirements have forced nuclear data researchers to examine closely the methods used in processing nuclear data. Important objectives are the elimination of bias in evaluations, and realistic assessment of uncertainties in the data base. The mathematical tools required to handle this problem have been available since the days of K. F. Gauss (ca. 1809), but they have been largely overlooked by the nuclear data community until recently. F. G. Perey of Oak Ridge National Laboratory has been a leader in pointing out the value of covariance and least-squares methods in nuclear data analyses. Although these methods are now generally known by nuclear data evaluators, they are apparently not well understood by many other nuclear data researchers. Evaluators are now requesting experimenters to provide more complete information on their experimental work, especially with regard to uncertainties and correlations. Therefore, it behoves experimentalists to acquire a basic understanding of covariances and least-squares methods so that an improvement in communication between experimentalists and evaluators can be achieved. Nuclear data experimentalists should not overlook an important fact of life: Experimental data rarely flows directly from the measurers to the people who use the results for nuclear energy applications. The results employed in applications are almost always evaluated results.

There appears to be a need to provide experimentalists and other non-evaluators in the nuclear data field with explicit guides to the subject. This report is intended to be one such guide. It is not a review of the subject, and no attempt has been made to provide an exhaustive list of relevant references. Furthermore, this report does not attempt to address all the aspects of covariance analysis which may be relevant to nuclear data research. The reader may wish to examine Refs. 1-6 since the present work was guided in part by these documents. References 7 and 8 are two other documents which were prepared with objectives similar to the present endeavor. Reference 7 is a fine explicit guide to covariance methods for the uninitiated. Reference 8 offers lucid insight into concepts of error analysis and data evaluation. Reference 9 is a small monograph which is clearly written and discusses numerous related topics although it does not address the present subject directly.

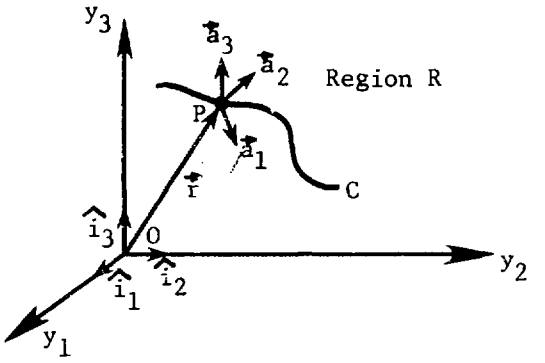
This report deals with the mechanics of error propagation, data evaluation and data fitting. The statistical aspects of error analysis will not be discussed. However, one important statistical point has to be emphasized: The errors assumed for all parameters considered in a particular analysis must conform to the same equivalent confidence level. For example, one cannot properly interpret the result of an error propagation calculation if the errors for some of the parameters are one standard deviation (1σ) and those for others are two standard deviations (2σ). Estimation of the confidence levels for errors in random variables which obey Normal statistics is straightforward. For variables which obey other statistical laws, or have systematic errors, the problem is more difficult. The practitioner must simply do the best he can to insure that the errors for all the considered parameters are "equivalent" in this respect.

Section II reviews those aspects of analytic geometry and vector calculus which provide insight into the mechanics of error analysis. It seems that this aspect has not previously received the emphasis it really deserves. Section III shows how to derive the error for a scalar function of several variables and Section IV extends the formalism to a vector function. Section V considers an application of the formalism in Section IV to a problem of special importance in the field of nuclear data. Section VI offers some comments on the preceding material, and it provides a natural stopping point for the reader who is interested in error propagation methods but not in evaluations. The basic method of least-squares evaluation is discussed in Section VII, with emphasis on the algebraic aspects of the theory. Section VIII provides the user with an explicit application of the methods discussed in Section VII. Techniques for dealing with non-linear evaluation problems are discussed in Section IX. Section X explains how an evaluator can update an evaluation to include new data sets without completely redoing the evaluation. This is an important feature of the least-squares method. Section XI is very important because it explains by means of examples how the concepts of Section VII and IX can be used by experimenters in everyday data analysis activities. Finally, Section XII offers some general comments on the subject material.

II. CURVILINEAR COORDINATES AND METRIC COEFFICIENTS - ORIGIN OF THE FORMALISM

Consider, for example, a region R of three-dimensional space. Often we describe points P in this region by giving their Cartesion coordinates (y_1, y_2, y_3) relative to some origin O.

But, we could define many other more complicated coordinate systems, say one where point P is given by (x_1, x_2, x_3).



Consider such an alternate coordinate system which is linked to our Cartesion set by the functional relations:

$$x_1 = x_1(y_1, y_2, y_3) \quad , \quad x_2 = x_2(y_1, y_2, y_3) \quad , \quad x_3 = x_3(y_1, y_2, y_3) \quad . \tag{1}$$

These relations must be single-valued and continuously differentiable over the region R, and it must be possible to obtain the inverse relations over the same region:

$$y_1 = y_1(x_1, x_2, x_3) \quad , \quad y_2 = y_2(x_1, x_2, x_3) \quad , \quad y_3 = y_3(x_1, x_2, x_3) \quad . \tag{2}$$

The relations of Eq. (1) define a transformation from Cartesian coordinates to general curvilinear coordinates, while Eq. (2) provides the inverse transformation.

Consider a curve C through region R. Let dL be a very small element of arc along the curve at point P. Let $\vec{r}$ be the vector from origin O to point P. We are permitted for a Cartesian coordinate system to write down the following expressions directly:

$$\vec{r} = \hat{i}_1 y_1 + \hat{i}_2 y_2 + \hat{i}_3 y_3 \quad , \tag{3}$$

$$d\vec{r} = \hat{i}_1 dy_1 + \hat{i}_2 dy_2 + \hat{i}_3 dy_3 \quad , \tag{4}$$

$$(dL)^2 = d\vec{r} \bullet d\vec{r} = \sum_{i=1}^{3} (dy_i)^2 \quad , \tag{5}$$

where $\hat{i}_1$, $\hat{i}_2$ and $\hat{i}_3$ are unit vectors along the y_1, y_2, and y_3 axes of the Cartesion coordinate system and $d\vec{r}$ is the small incremental change in $\vec{r}$ corresponding to changes dy_1, dy_2 and dy_3. The vector $d\vec{r}$ is tangent to the curve C and $(dL)^2$ is the square of its length. Note that "•" implies a vector dot product.

Matters become more complicated when we consider the curve C in curvilinear coordinates defined by x_1, x_2 and x_3. Then

$$d\vec{r} = \frac{\partial \vec{r}}{\partial x_1} dx_1 + \frac{\partial \vec{r}}{\partial x_2} dx_2 + \frac{\partial \vec{r}}{\partial x_3} dx_3 \tag{6}$$

where

$$\frac{\partial \vec{r}}{\partial x_i} = \lim_{x_i \to 0} \frac{\Delta \vec{r}}{\Delta x_i} \quad (x_j \text{ for } j \neq i \text{ held fixed}). \tag{7}$$

For simplicity, define the <u>base vectors</u> $\vec{a}_i$

$$\vec{a}_i = \frac{\partial \vec{r}}{\partial x_i} \; (i = 1, 2, 3) \quad , \tag{8}$$

then

$$d\vec{r} = \vec{a}_1 \, dx_1 + \vec{a}_2 \, dx_2 + \vec{a}_3 \, dx_3 \quad . \tag{9}$$

Vector equations apply regardless of the choice of coordinate system, so Eq. (5) can be express in curvilinear coordinates by

$$(dL)^2 = d\vec{r} \bullet d\vec{r} = \sum_{i=1}^{3} \sum_{j=1}^{3} (\vec{a}_i \bullet \vec{a}_j) dx_i \, dx_j \tag{10}$$

Let G_{ij} be defined as

$$G_{ij} = \vec{a}_i \bullet \vec{a}_j \quad , \tag{11}$$

then

$$(dL)^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} G_{ij} \, dx_i \, dx_j \quad . \tag{12}$$

Equation (12) is an example of a quadratic differential form, and it is clearly symmetric since

$$G_{ij} = G_{ji} \quad . \tag{13}$$

The coefficients G_{ij} are called the metric coefficients for this particular curvilinear coordinate system at point P. A Cartesian coordinate system is a special case of a curvilinear coordinate system with

$$G_{ij} = \delta_{ij} \quad , \tag{14}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad . \tag{15}$$

The quantity δ_{ij} is the Kronecker Delta. The elements G_{ij} form a matrix $\overline{G}$ where, explicitly,

$$\overline{G} = \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{bmatrix} \quad . \tag{16}$$

The quadratic differential form from Eq. (12) can be expressed in matrix notation,

$$(dL)^2 = (d\vec{x})^T \bullet \overline{G} \bullet (d\vec{x}) \quad , \tag{17}$$

or more explicitly (to help visualize the analysis):

$$(dL)^2 = \left(dx_1, dx_2, dx_3\right) \bullet \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{bmatrix} \bullet \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} \quad . \tag{18}$$

Here, "$\bullet$" implies matrix multiplication and "T" designates a transposed matrix (column and row elements interchanged).

We can alter the representation somewhat by defining

$$\vec{a}_i = s_i \hat{u}_i \qquad (i = 1, 2, 3) \tag{19}$$

where $\hat{u}_i$ is a unit vector parallel to the base vector $\vec{a}_i$ which points in the direction corresponding to an increase in x_i. Thus, Eq. (10) becomes

$$(dL)^2 = d\vec{r} \bullet d\vec{r} = \sum_{i=1}^{3} \sum_{j=1}^{3} (\hat{u}_i \bullet \hat{u}_j) s_i s_j \, dx_i dx_j \quad . \tag{20}$$

Define

$$C_{ij} = \hat{u}_i \bullet \hat{u}_j \quad , \tag{21}$$

then

$$(dL)^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} C_{ij} \, s_i s_j dx_i dx_j \quad . \tag{22}$$

If $\bar{C}$ is defined as the matrix of elements C_{ij}, then

$$(dL)^2 = (\bar{S} \bullet d\vec{x})^T \bullet \bar{C} \bullet (\bar{S} \bullet d\vec{x}) \tag{23}$$

in matrix notation.

$\bar{S}$ will be referred to as the <u>sensitivity matrix</u>, and $\bar{C}$ as the <u>correlation matrix</u> for this curvilinear coordinate system. The matrix $\bar{S}$ is a diagonal matrix which is given explicitly by

$$\bar{S} = \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{bmatrix} \quad . \tag{24}$$

The metric coefficients G_{ij} can be calculated readily if the inverse transformation relations from Eq. (2) are known. We have

$$dy_k = \sum_{i=1}^{3} \left(\frac{\partial y_k}{\partial x_i}\right) dx_i \qquad (k = 1, 2, 3) \quad , \tag{25}$$

then, from a comparison of Eq. (5) and Eq. (12), one can deduce the expression

$$G_{ij} = \sum_{k=1}^{3} \left(\frac{\partial y_k}{\partial x_i}\right)\left(\frac{\partial y_k}{\partial x_j}\right) \qquad (i,j = 1, 2, 3) \quad , \tag{26}$$

following some routine algebra.

What has all of this to do with errors? To answer this question, let us summarize the preceding material in words: We consider a three-dimensional region of space. The invariant quantity to be examined is the distance along some curve in the space. When we look at a small local element of length near a particular point, we find that it can be calculated in terms of small increments of the curvilinear coordinates provided that one knows the matrix of metric coefficients, or alternatively, the sensitivity matrix and the correlation matrix. The calculation is performed using the quadratic differential form. The analogy of this formalism to the problem of error propagation should begin to be evident. The parameters of the experiment play the role of the curvilinear coordinates. Points along the curve correspond to various experimental outcomes for various values of the parameters. The "distance" between two nearby points, which differ only due to small positive or negative increments in the coordinates (errors in the parameters), corresponds to the uncertainty in the experimental result. This uncertainty is to be calculated using the quadratic differential form, and one must have a knowledge of the sensitivity of the experimental result to the experimental parameters (sensitivity matrix) and the correlations between the parameters (correlation matrix). The application of these basic concepts to error analysis will be pursued in the following sections of this report.

Before going on to the next section, there are two items of unfinished business to complete. One does not normally speak of a covariance matrix in discussions of the metrics of curvilinear coordinates, however, we will now define a matrix $\overline{M}$ which plays the role of the <u>covariance matrix</u> when this formalism is applied to error analysis. Refer to Eq. (22), then consider a matrix $\overline{M}$ which consists of the elements

$$M_{ij} = C_{ij}\, dx_i\, dx_j \qquad (i,j = 1, 2, 3) \quad . \tag{27}$$

Define $\vec{I}$ as follows:

$$\vec{I} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad . \tag{28}$$

Then Eq. (22) can be written in the form

$$(dL)^2 = (\bar{S} \bullet \vec{I})^T \bullet \bar{M} \bullet (\bar{S} \bullet \vec{I}) \quad . \tag{29}$$

What is the meaning of the correlation matrix $\bar{C}$? The elements C_{ij} satisfy the condition

$$-1 \leq C_{ij} \leq 1 \quad . \tag{30}$$

The diagonal elements C_{ii} are always unity. For Cartesian coordinate systems, the off-diagonal elements are always zero. A Cartesian coordinate system is an <u>orthogonal</u> system. The variables y_1, y_2, and y_3 can vary independently of one another. Curvilinear coordinate systems may be orthogonal, but they don't have to be. If the base vectors $\vec{a}_i$ at a particular point P are non-orthogonal, then they are not all perpendicular. A small change in position involving a change dx_i will involve an unavoidable change in the value of another x_j for which the coefficient C_{ij} is non-zero. When the formalism is applied to error analysis, one sees that the errors in any two parameters are correlated whenever such a non-orthogonal relationship exists locally between the two parameters.

III. ERROR ANALYSIS FOR A SCALAR QUANTITY

Suppose we wish to perform an experiment to determine a physical quantity V which is a function of several (say n) experimental parameters $x_1, \ldots, x_n$. It will be assumed that the relationship between the variables x_i and the derived quantity V is smooth. Using the language of mathematics, we expect continuity and differentiability of the function defined by

$$V = V(x_1, \ldots, x_n) = V(\vec{x}) \quad . \tag{31}$$

Let E_V be the uncertainty (error) in V which results from the uncertainties E_{x1}, $\ldots$, E_{xn} ($\vec{E}_x$) in the parameters $\vec{x}$. After having been exposed to the material in the preceding section, the reader will not be surprised to learn that

$$E_V{}^2 = (\bar{S} \bullet \vec{E}_x)^T \bullet \bar{C}_x \bullet (\bar{S} \bullet \vec{E}_x) \tag{32}$$

is the formula to apply in deriving the error in V, with $\bar{C}_x$ defined as the matrix of correlations between the errors $E_{x1}, \ldots, E_{xn}$, and

$$\bar{S} = \begin{bmatrix} (\partial V/\partial x_1) & & 0 \\ & \ddots & \\ 0 & & (\partial V/\partial x_n) \end{bmatrix} \quad . \tag{33}$$

The matrices are of apriori size n x n, but analysis can proceed with smaller-size matrices when circumstances produce redundancy for the full n x n matrices.

The information in the <u>error vector</u> $\vec{E}_x$ and <u>correlation matrix</u> $\bar{C}_x$ can be combined into the <u>covariance</u> matrix $\bar{M}_x$, defined in terms of its elements by

$$M_{xij} = C_{xij}\, E_{xi}\, E_{xj} \qquad (i,j = 1,n) \quad , \tag{34}$$

$$E_{xi} = (M_{xii})^{1/2} \qquad (i = 1,n) \quad . \tag{35}$$

Then the equivalent to Eq. (32) is

$$E_V{}^2 = (\bar{S} \bullet \vec{I})^T \bullet \bar{M}_x \bullet (\bar{S} \bullet \vec{I}) \quad , \tag{36}$$

where $\vec{I}$ is a vector like the one defined in Eq. (28), with dimension n.

It is often preferable to work with $\vec{E}_x$ and $\overline{C}_x$ rather than with $\overline{M}_x$ because the elements of $\vec{E}_x$ give the explicit errors in the parameters while $\overline{C}_x$ gives an indication of the degree of correlation between these various errors - in dimensionless form. $\overline{M}_x$ is harder to understand.

Consider the vector differential $d\vec{V}$ defined by

$$d\vec{V} = \sum_{i=1}^{n} \left(\frac{\partial V}{\partial x_i}\right) d\vec{x}_i \quad . \tag{37}$$

Then, if the symbol <...> is used to indicate an <u>expected value</u>,

$$\langle d\vec{V} \bullet d\vec{V}\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial V}{\partial x_i}\right) \left(\frac{\partial V}{\partial x_j}\right) \langle d\vec{x}_i \bullet d\vec{x}_j\rangle = E_V{}^2 \quad , \tag{38}$$

from consideration of Eq. (36), provided that

$$M_{xij} = \langle d\vec{x}_i \bullet d\vec{x}_j\rangle \tag{39}$$

Equation (39) is often used as the definition of the covariance matrix. The larger the assumed errors, the smaller the probability that one will observe deviations which exceed these errors. For example, the probability of a deviation exceeding three standard deviations (3σ) is only about 0.3% for Normal statistics.

How should one go about deriving the covariance matrix for an experiment? The first step is to catalogue all the <u>sources</u> of error in the established parameters $\vec{x}$ of the experiment and decide what sort of correlations exist. The magnitudes of the various error components should be determined and these errors should be expressed in the <u>units of the corresponding parameters.</u> The result of this exercise would be a table of the form:

		Error Components						
	$i \backslash \ell$	1	2	...	ℓ'	...	L	Total Error
Parameter	1	e_{11}	e_{12}	...	$e_{1\ell'}$	...	e_{1L}	E_{x_1}
	2	e_{21}	e_{22}	...	$e_{2\ell'}$	...	e_{2L}	E_{x_2}
	⋮	⋮			⋮		⋮	⋮
	i'	$e_{i'1}$	$e_{i'2}$	...	$e_{i'\ell'}$	...	$e_{i'L}$	$E_{xi'}$
	⋮	⋮	⋮		⋮		⋮	⋮
	n	e_{n1}	e_{n2}	...	$e_{n\ell'}$	...	e_{nL}	E_{xn}

The quantity $e_{i\ell}$ is the <u>magnitude</u> of the uncertainty in x_i due to effect "ℓ". We assume that a total of L different such effects are considered. Clearly some $e_{i\ell}$ may be zero. The elements of the covariance matrix $\overline{M}$ are calculated from the expression

$$M_{xij} = \sum_{\ell = 1}^{L} S_{ij\ell}\, e_{i\ell}\, e_{j\ell} \quad (i,\ j = 1,\ n) \quad . \tag{40}$$

The $S_{ij\ell}$ are correlation parameters relating error components for x_i and x_j. They have the following properties:

$$S_{ii\ell} = 1 \quad (i = 1,\ n \text{ and } \ell = 1,\ L) \quad , \tag{41}$$

$$-1 \leq S_{ij\ell} \leq +1 \quad (i \neq j, \text{ and } \ell = 1,\ L) \quad . \tag{42}$$

When the errors in x_i and x_j due to effect "ℓ" are totally uncorrelated, then $S_{ij\ell}$ must be zero. When the errors in x_i and x_j due to effect "ℓ" are totally correlated in the positive sense, then $S_{ij\ell}$ must be + 1. Total anticorrelation implies a value − 1. Partial correlation or anti-correlation leads to intermediate values of $S_{ij\ell}$. Since the experimenter does not usually derive the covariance matrix, he should provide an error component table and enough information about the experiment to enable someone else (e.g. an evaluator) to determine the $S_{ij\ell}$ values and generate the covariance matrix.

Experimenters should try as much as possible to establish the various effects "ℓ" to be considered so that the $S_{ij\ell}$ will be either − 1 (rarely), 0, or + 1. The experimenter is on rather uncertain ground if he is in a position of having to estimate partial correlations or anti-correlations.

Example: Error in the ratio of two cross sections

Two cross sections σ_1 and σ_2 are available and one wishes to evaluate the error in the ratio

$$R = R(\sigma_1, \sigma_2) = \sigma_1/\sigma_2 \quad , \tag{43}$$

if each cross section has a 1% uncorrelated error component and a 2% fully correlated error component. The correlation is assumed to arise because each cross section was measured using the same standard; an alteration in this standard would cause σ_1, and σ_2 to change in the same direction. They are thus correlated in the positive sense.

The first step toward generation of the covariance matrix for the errors in cross sections σ_1 and σ_2 is to construct a table which exhibits the correlated, uncorrelated and total errors. This table is the basis for calculation of the covariance matrix elements:

Parameter	Uncorrelated Error	Correlated Error	Total Error
σ_1	$0.01\ \sigma_1$	$0.02\ \sigma_1$	$[(0.01\ \sigma_1)^2 + (0.02\ \sigma_1)^2]^{1/2}$
σ_2	$0.01\ \sigma_2$	$0.02\ \sigma_2$	$[(0.01\ \sigma_2)^2 + (0.02\ \sigma_2)^2]^{1/2}$

Define $E_{\sigma 1}$ and $E_{\sigma 2}$ to be the total errors in σ_1 and σ_2 respectively, $\overline{M}$ to be the covariance matrix, $\overline{S}$ to be the sensitivity matrix and E_R to be the error in the ratio R. The sensitivity matrix is given explicitly by

$$\overline{S} = \begin{bmatrix} (\partial R/\partial \sigma_1 & 0 \\ 0 & (\partial R/\partial \sigma_2) \end{bmatrix} = \begin{bmatrix} 1/\sigma_2 & 0 \\ 0 & -\sigma_1/\sigma_2^2 \end{bmatrix} \quad , \tag{44}$$

and the covariance matrix elements, according to Eq. (40), are:

$$M_{11} = E_{\sigma 1}{}^2 = (0.01\ \sigma_1)^2 + (0.02\ \sigma_1)^2$$

$$M_{12} = M_{12} = +\ (0.02\ \sigma_1)\ (0.02\ \sigma_2) \tag{45}$$

$$M_{22} = E_{\sigma_2}{}^2 = (0.01\ \sigma_2)^2 + (0.02\ \sigma_2)^2 \quad .$$

The off-diagonal element M_{12} is positive and is the product of correlated errors for the two cross sections σ_1 and σ_2. From Eqs. (34) and (35), it is readily seen that the off-diagonal element C_{12} of the correlation matrix $\bar{C}$ is

$$C_{12} = C_{21} = M_{12}/(E_{\sigma 1}\ E_{\sigma 2}) = 0.8 \quad . \tag{46}$$

The error E_R in the ratio R is then calculated from

$$E_R{}^2 = (\bar{S} \bullet \vec{E})^T \bullet \bar{C} \ \bullet (\bar{S} \bullet \vec{E}) \quad , \tag{47}$$

with

$$\vec{E} = \begin{bmatrix} E_{\sigma 1} \\ \\ E_{\sigma 2} \end{bmatrix} \quad . \tag{48}$$

Straightforward matrix multiplication yields the result

$$E_R^2 = \left(\frac{1}{\sigma_2}\right)^2 E_{\sigma 1}{}^2 + \left(\frac{-\sigma_1}{\sigma_2{}^2}\right)^2 E_{\sigma 2}{}^2 + 2(0.8)\left(\frac{1}{\sigma_2}\right)\left(\frac{-\sigma_1}{\sigma_2{}^2}\right) E_{\sigma 1}\ E_{\sigma 2} \tag{49}$$

$$\left(\frac{E_R}{R}\right)^2 = \left(\frac{E_{\sigma 1}}{\sigma_1}\right)^2 + \left(\frac{E_{\sigma 2}}{\sigma_2}\right)^2 - 2(0.8)\left(\frac{E_{\sigma 1}}{\sigma_1}\right)\left(\frac{E_{\sigma 2}}{\sigma_2}\right) \quad . \tag{50}$$

Since $(E_{\sigma 1}/\sigma_1) = (E_{\sigma 2}/\sigma_2) \approx 0.02236$, we obtain $(E_R/R) \approx 0.01414$. If the correlation had been ignored by eliminating the last term from Eq. (49), then $(E_R/R) \approx 0.03162$ would have been obtained. In this example, a neglect of the correlation leads to the determination of too large an error for the ratio R.

For this problem, the correlation leads to a smaller error because the ratio is a quotient of the two cross sections σ_1 and σ_2, and one element of the sensitivity matrix is negative. If the error in the product of the cross sections had been sought, then both of the elements of the sensitivity matrix would have been positive and the sign of the last term in Eq. (49) or Eq. (50) would have been positive. This would mean that inclusion of the correlation would increase the error.

The expression for M_{12} in Eq. (45) is explicitly positive since the definition of the problem indicates that the correlated portions of the errors in σ_1 and σ_2 are positively correlated (e.g., if one is measured too large, then so is the other). There can be instances where two parameters are anti-correlated (e.g., when one increases the other decreases). If this is the case, a negative sign must be inserted in front of the corresponding covariance matrix element, and the correlation coefficient will then be negative as allowed by Eq. (30). So, one must be careful to distinguish the effects due to the sign of the correlation of the parameters and those due to the sign of the sensitivity matrix elements. The correlation parameters $S_{ij\ell}$ appearing in Eq. (40) provide us with a systematic way to keep track of these considerations.

IV. ERROR ANALYSIS FOR A VECTOR QUANTITY

The formalism in Sec. III is adequate to handle the error analysis for a single quantity derived from several experimental parameters. However, one often wishes to analyze the errors for several quantities derived from a set of experimental parameters. It is desired to deduce the correlations between the errors in the derived quantities as well as the magnitudes of their errors. To do this, it is necessary to use a vector formalism.

Assume that m quantities $V_1, \ldots, V_m$ $(\vec{V})$ are to be derived from n parameters $x_1, \ldots, x_n$ $(\vec{x})$. The relationships are defined by m continuous, differentiable functions represented by the vector equation

$$\vec{V} = \vec{V}(\vec{x}) \quad . \tag{51}$$

The errors in the parameters $\vec{x}$ are represented by the vector $\vec{E}_x$, and $\bar{C}_x$ is the correlation matrix for these errors. Then the covariance matrix $\bar{M}_V$ for the derived quantities is given in terms of its elements by

$$M_{Vij} = \left(\bar{S}_i \bullet \vec{E}_x\right)^T \bullet \bar{C}_x \bullet \left(\bar{S}_j \bullet \vec{E}_x\right) \quad (i,j = 1,m) \quad . \tag{52}$$

$\bar{M}_V$ has size m x m since $\vec{V}$ has size m. $\bar{C}_x$ has size n x n and $\vec{E}_x$ has size n since $\vec{x}$ has size n. All the $\bar{S}_i$ have size n x n, and these various sensitivity matrices have the explicit form

$$\bar{S}_i = \begin{bmatrix} (\partial V_i/\partial x_1) & & & & 0 \\ & \ddots & & & \\ & & (\partial V_i/\partial x_k) & & \\ & & & \ddots & \\ 0 & & & & (\partial V_i/\partial x_n) \end{bmatrix} \quad (i = 1, m) \quad . \tag{53}$$

The error vector $\vec{E}_V$ for $\vec{V}$ is derived from

$$E_{Vi} = (M_{Vii})^{1/2} \quad (i = 1, m) \quad , \tag{54}$$

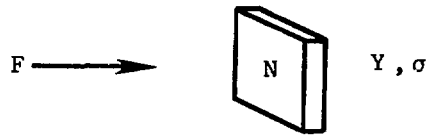
and the correlation matrix $\bar{C}_V$ for the errors in the derived quantities is deduced from

$$C_{Vij} = M_{Vij}/(E_{Vi}\, E_{Vj}) \quad (i, j = 1, m) \quad . \tag{55}$$

Clearly, $\vec{E}_V$ has size m and $\bar{C}_V$ has size m x m.

Example: Errors and correlations for a cross section set

Consider two foils of different material which are exposed to the same neutron flux. The induced activities are measured with the same detector and cross sections are calculated using the formulas



$$\sigma_1 = Y_1/(FN_1) \quad , \qquad (56)$$

$$\sigma_2 = Y_2/(FN_2) \quad . \qquad (57)$$

Y = yield
N = atoms
F = neutron flux

The uncorrelated and correlated errors are given in the following table:

	Error in Y		Error in N		Error
Foil	Uncorrelated	Correlated	Uncorrelated	Correlated	in F
1	3%	2%	0.5%	1%	4%
2	5%	2%	0.7%	1%	4%

The error correlations implied by this table are between Y_1 and Y_2, and between N_1 and N_2; the error in the flux F is fully correlated in the two measurements because both cross sections were measured in the same flux. We can assume that the error correlation for Y_1 and Y_2 arises from a common detector efficiency calibration factor, while the error correlation for N_1 and N_2 arises because the foils were both weighed with the same balance.

The formalism demands that five experimental parameters be considered. Therefore, we write

$$\sigma_1 = \sigma_1 \, (Y_1, Y_2, N_1, N_2, F) \quad , \qquad (58)$$

$$\sigma_2 = \sigma_2 \, (Y_1, Y_2, N_1, N_2, F) \quad . \qquad (59)$$

We will see that the fact that Y_2 and N_2 are dummy variables in Eq. (58) while Y_1 and N_1 are dummy variables in Eq. (59) leads to some simplification at a later stage of the analysis. We can write

$$\vec{x} = \begin{bmatrix} Y_1 \\ Y_2 \\ N_1 \\ N_2 \\ F \end{bmatrix} = \begin{bmatrix} \vec{Y} \\ \vec{N} \\ F \end{bmatrix} \qquad (60)$$

and

$$\vec{E}_x = \begin{bmatrix} E_{Y1} \\ E_{Y2} \\ E_{N1} \\ E_{N2} \\ E_F \end{bmatrix} = \begin{bmatrix} \vec{E}_Y \\ \vec{E}_N \\ E_F \end{bmatrix} \tag{61}$$

for the parameter vector and vector of parameter errors respectively. Use is made of the notation of sub-vectors of a vector: Analogously, we can express the covariance matrix $\overline{M}_x$ in terms of submatrices:

$$\overline{M}_x = \begin{bmatrix} \overline{M}_{YY} & \overline{M}_{YN} & \overline{M}_{YF} \\ \overline{M}_{NY} & \overline{M}_{NN} & \overline{M}_{NF} \\ \overline{M}_{FY} & \overline{M}_{FN} & \overline{M}_{FF} \end{bmatrix} . \tag{62}$$

Since F, E_F and M_{FF} are single elements, they are not affixed with vector or matrix lables. The formalism demands that correlations between all five experimental parameters be considered. From the nature of the problem, it is clear that all the off-diagonal submatrices in Eq. (62) are zero matrices (all elements zero) since no correlations exist between the different categories of variables Y, N and F. Therefore,

$$\overline{M}_x = \begin{bmatrix} \overline{M}_{YY} & & 0 \\ & \overline{M}_{NN} & \\ 0 & & M_{FF} \end{bmatrix} . \tag{63}$$

Likewise, for the correlation matrix $\overline{C}_x$ we have

$$\overline{C}_x = \begin{bmatrix} \overline{C}_{YY} & & 0 \\ & \overline{C}_{NN} & \\ 0 & & 1 \end{bmatrix} . \tag{64}$$

Two sensitivity matrices of size 5 x 5 must be considered:

$$\bar{S}_1 = \begin{bmatrix} (\partial\sigma_1/\partial Y_1) & 0 & & & 0 \\ 0 & 0 & & & \\ & & (\partial\sigma_1/\partial N_1) & 0 & \\ & & 0 & 0 & \\ 0 & & & & (\partial\sigma_1/\partial F) \end{bmatrix} = \begin{bmatrix} \bar{S}_{1Y} & & 0 \\ & \bar{S}_{1N} & \\ 0 & & S_{1F} \end{bmatrix} \quad (65)$$

and

$$\bar{S}_2 = \begin{bmatrix} 0 & 0 & & & 0 \\ 0 & (\partial\sigma_2/\partial Y_2) & & & \\ & & 0 & 0 & \\ & & 0 & (\partial\sigma_2/\partial N_2) & \\ 0 & & & & (\partial\sigma_2/\partial F) \end{bmatrix} = \begin{bmatrix} \bar{S}_{2Y} & & 0 \\ & \bar{S}_{2N} & \\ 0 & & S_{2F} \end{bmatrix}. \quad (66)$$

The fact that $(\partial\sigma_1/\partial Y_2)$, $(\partial\sigma_1/\partial N_2)$, $(\partial\sigma_2/\partial Y_1)$ and $(\partial\sigma_2/\partial N_1)$ are zero is reflected in Eqs. (65) and (66). The remaining nonzero partial derivatives can be calculated using the formulas:

$$(\partial\sigma_i/\partial Y_i) = 1/(FN_i) = \sigma_i/Y_i \qquad (i = 1, 2) \quad , \qquad (67)$$

$$(\partial\sigma_i/\partial N_i) = -Y_i/(FN_i^2) = -\sigma_i/N_i \qquad (i = 1, 2) \quad , \qquad (68)$$

$$S_{iF} = (\partial\sigma_i/\partial F) = -Y_i/(N_iF^2) = -\sigma_i/F \qquad (i = 1, 2) \quad . \qquad (69)$$

The covariance matrix elements are derived using the expression

$$M_{\sigma ij} = (\bar{S}_i \bullet \vec{E}_x)^T \bullet \bar{C}_x \bullet (\bar{S}_j \bullet \vec{E}_x) \qquad (i, j = 1, 2) \quad . \qquad (70)$$

It is left as an exercise for the reader to show that an application of matrix algebra leads to

$$M_{\sigma ij} = (\bar{S}_{iY} \bullet \vec{E}_Y)^T \bullet \bar{C}_{YY} \bullet (\bar{S}_{jY} \bullet \vec{E}_Y)$$

$$+ (\bar{S}_{iN} \bullet \vec{E}_N)^T \bullet \bar{C}_{NN} \bullet (\bar{S}_{jN} \bullet \vec{E}_N) \tag{71}$$

$$+ S_{iF}\, S_{jF}\, E_F^{\,2} \qquad (i,\ j = 1,\ 2) \quad .$$

For this example, a size -5 problem reduces to two size -2 problems plus a size -1 problem owing to the absence of correlations between several of the parameters. Reductions in size of this sort happen often in covariance analyses. The practitioner of the method should always look for such "short cuts" to reduce labor. Further matrix algebra leads to the result:

$$M_{\sigma 11} = \sigma_1^{\,2}\left[\left(\frac{E_{Y1}}{Y_1}\right)^2 + \left(\frac{E_{N1}}{N_1}\right)^2 + \left(\frac{E_F}{F}\right)^2\right] \quad , \tag{72}$$

$$M_{\sigma 12} = M_{\sigma 21} = \sigma_1\sigma_2\left[C_{YY12}\left(\frac{E_{Y1}}{Y_1}\right)\left(\frac{E_{Y2}}{Y_2}\right) + C_{NN12}\left(\frac{E_{N1}}{N_1}\right)\left(\frac{E_{N2}}{N_2}\right) + \left(\frac{E_F}{F}\right)^2\right] \quad , \tag{73}$$

$$M_{\sigma 22} = \sigma_2^{\,2}\left[\left(\frac{E_{Y2}}{Y_2}\right)^2 + \left(\frac{E_{N2}}{N_2}\right)^2 + \left(\frac{E_F}{F}\right)^2\right] \quad . \tag{74}$$

The components of $\vec{E}_x$ are deduced from data in the table as follows:

$$\vec{E}_Y \begin{cases} E_{Y1}^{\,2} = (0.03\ Y_1)2 + (0.02\ Y_1)2 \quad , & (75) \\ E_{Y2}^{\,2} = (0.05\ \ Y_2)2 + (0.02\ Y_2)2 \quad , & (76) \end{cases}$$

$$\vec{E}_N \begin{cases} E_{N1}^{\,2} = (0.005\ N_1)2 + (0.01\ N_1)2 \quad , & (77) \\ E_{N2}^{\,2} = (0.007\ N_2)^2 + (0.01\ N_2)^2 \quad , & (78) \end{cases}$$

$$E_F^{\,2} = M_{FF} = (0.04\ F)^2 \quad . \tag{79}$$

In order to evaluate $M_{\sigma 12}$ according to Eq. (73), C_{YY12} and C_{NN12} must be determined. To accomplish this, it is necessary to generate the covariance matrixes $\overline{M}_{YY}$ and $\overline{M}_{NN}$. The covariance matrix M_{YY} has the explicit form

$$\overline{M}_{YY} = \begin{bmatrix} E_{Y1}{}^2 & (0.02\ Y_1)\ (0.02\ Y_2) \\ (0.02\ Y_1)\ (0.02\ Y_2) & E_{Y2}{}^2 \end{bmatrix} , \tag{80}$$

and the covariance matrix $\overline{M}_{NN}$ has the explicit form

$$\overline{M}_{NN} = \begin{bmatrix} E_{N1}{}^2 & (0.01\ N_1)\ (0.01\ N_2) \\ (0.01\ N_1)\ (0.01\ N_2) & E_{N2}{}^2 \end{bmatrix} . \tag{81}$$

The off-diagonal elements of $\overline{M}_{YY}$ and $\overline{M}_{NN}$ are both positive, reflecting the positive nature of the correlations involved (as opposed to anti-correlations). The off-diagonal element of $\overline{C}_{YY}$ is given by

$$C_{YY12} = C_{YY21} = M_{YY12}/(E_{Y1}\ E_{Y2}) \approx 0.2060 \quad , \tag{82}$$

and the off-diagonal element of $\overline{C}_{NN}$ is given by

$$C_{NN12} = C_{NN21} = M_{NN12}/(E_{N1}\ E_{N2}) \approx 0.7327 \quad . \tag{83}$$

Following more algebra, we learn that the fractional errors in σ_1 and σ_2 are:

$$(E_{\sigma 1}/\sigma_1) = (M_{\sigma 11})^{1/2}/\sigma_1 = 0.055 \quad , \tag{84}$$

$$(E_{\sigma 2}/\sigma_2) = (M_{\sigma 22})^{1/2}/\sigma_2 \approx 0.06818 \quad . \tag{85}$$

The correlation coefficient linking these errors is

$$C_{\sigma 12} = C_{\sigma 21} = M_{\sigma 12}/(E_{\sigma 1}E_{\sigma 2}) \approx 0.5600 \quad . \tag{86}$$

It is clear from Eqs. (72) and (74) that we could have deduced the fractional errors in σ_1 and σ_2 by simply combining the errors in the corresponding experimental parameters in quadrature as one usually does. However, the preceding detailed covariance analysis was required to show that there is a substantial correlation (~56%) between these errors. It is evident from this simple example that covariance analysis is a lot of work. In realistic applications, it is often impractical to perform the analysis by hand and one must resort to the use of a digital computer. The preceding development started from first principles, as embodied in Eq. (70), and indicated most of the ensuing steps. In practice, one finds short cuts in most categories of problems, such as we did in the present example, which help to reduce the number of computational steps. It is beyond the scope of this report to consider all these possibilities.

V. COMBINATION OF RATIO AND STANDARD ERRORS

One type of problem which arises very often in the field of nuclear data is the following: One measures a set of ratios of some unknown (or lesser known) cross sections to a set of standard cross sections. The errors and correlations for the ratios can be deduced from a knowledge of the experimental parameters as described in the preceding section. Covariance information for the standard cross sections may be available from evaluations such as the ENDF/B-V Standards File (Ref. 10). The question is how to combine this information so as to obtain the covariance matrix for the unknown cross sections. The problem is addressed in this section.

Assume that n ratios R_i are measured, and that the corresponding standard cross sections are S_i. The unknown cross sections σ_i are derived from the relationship

$$\sigma_i = R_i \, S_i \qquad (i = 1, n) \quad . \tag{87}$$

In vector notation we have

$$\vec{\sigma} = \vec{\sigma} \, (\vec{R}, \vec{S}) \quad , \tag{88}$$

where $\vec{\sigma}$, $\vec{R}$ and $\vec{S}$ are vectors representing the unknown cross sections, the ratios, and the standard cross sections respectively. Assume that the ratio errors $\vec{E}_R$ and the ratio correlation matrix $\overline{C}_R$, as well as the standard errors $\vec{E}_S$ and the standard correlation matrix $\overline{C}_S$, are all available. Define

$$\vec{x} = \begin{bmatrix} \vec{R} \\ \vec{S} \end{bmatrix} \quad , \tag{89}$$

and

$$\overline{C}_x = \begin{bmatrix} \overline{C}_R & \overline{0} \\ \overline{0} & \overline{C}_S \end{bmatrix} \quad . \tag{90}$$

It is assumed, apriori, that no correlations exist between the ratio and the standard values, so the off-diagonal submatrices of $\overline{C}_x$ are set to zero.

The rule for error propagation states that the covariance matrix $\overline{M}_\sigma$ for the unknown cross section can be expressed in terms of its elements by

$$M_{\sigma ij} = (\overline{S}_i \bullet \vec{E}_x)^T \bullet \overline{C}_x \bullet (\overline{S}_j \bullet \vec{E}_x) \qquad (i,\ j = 1,\ n) \quad , \tag{91}$$

where

$$\overline{S}_i = \begin{bmatrix} \overline{S}_{Ri} & \overline{0} \\ \overline{0} & \overline{S}_{Si} \end{bmatrix} \qquad (i = 1,\ n) \quad , \tag{92}$$

$$\overline{S}_{Ri} = \begin{bmatrix} \ddots & & 0 \\ & (\partial\sigma_i/\partial R_i) & \\ 0 & & \ddots \end{bmatrix} \qquad (i = 1,\ n) \quad , \tag{93}$$

$$\overline{S}_{Si} = \begin{bmatrix} \ddots & & 0 \\ & (\partial\sigma_i/\partial S_i) & \\ 0 & & \ddots \end{bmatrix} \qquad (i = 1,\ n) \quad . \tag{94}$$

Each of the matrices $\overline{S}_{Ri}$ and $\overline{S}_{Si}$ has <u>only one</u> nonzero element, located in the ith position on the diagonal. Clearly, from Eq. (87)

$$(\partial\sigma_i/\partial R_i) = S_i = \sigma_i/R_i \qquad (i = 1,\ n) \quad , \tag{95}$$

$$(\partial\sigma_i/\partial S_i) = R_i = \sigma_i/S_i \qquad (i = 1,\ n) \quad . \tag{96}$$

Straightforward matrix manipulations lead to the result

$$M_{\sigma ij} = (\bar{S}_{Ri} \bullet \vec{E}_R)^T \bullet \bar{C}_R \bullet (\bar{S}_{Rj} \bullet \vec{E}_R)$$

$$+ (\bar{S}_{Si} \bullet \vec{E}_S)^T \bullet \bar{C}_S \bullet (\bar{S}_{Sj} \bullet \vec{E}_S) \qquad (i,\ j = 1,\ n) \quad . \tag{97}$$

Once again, we see that an absence of correlations between two sets of parameters in a problem leads to a separation of the analysis into distinct parts, and reduces the size of the problem. Further analysis leads to the relatively simple result

$$M_{\sigma ij} = \sigma_i \sigma_j \left[C_{Rij} \left(\frac{E_{Ri}}{R_i}\right)\left(\frac{E_{Rj}}{R_j}\right) + C_{Sij} \left(\frac{E_{Si}}{S_i}\right)\left(\frac{E_{Sj}}{S_j}\right)\right] \qquad (i,\ j = 1,\ n) \quad . \tag{98}$$

From this expression, one can readily calculate the unknown cross section errors $\vec{E}_\sigma$ and the corresponding correlation matrix $\bar{C}_\sigma$ using the expressions

$$E_{\sigma i} = (M_{\sigma ii})^{1/2} = \sigma_i \left[\left(\frac{E_{Ri}}{R_i}\right)^2 + \left(\frac{E_{Si}}{S_i}\right)^2\right]^{1/2} \qquad (i = 1,\ n) \quad , \tag{99}$$

$$C_{\sigma ij} = M_{\sigma ij}/(E_{\sigma i}\ E_{\sigma j}) \qquad (i,\ j = 1,\ n) \quad . \tag{100}$$

VI. SOME COMMENTS RELEVANT TO THE GENERATION OF - AND ERROR PROPAGATION WITH - COVARIANCE MATRICES

If the reader has mastered the preceding sections, he will now understand how covariance matrices can be generated from properly specified experimental errors, and he will know how to use the covariance formalism in error analyses. At this point, it seems worthwhile to offer some suggestions, and to indicate the current thinking within the nuclear data evaluation community on this subject.

The ENDF/B-V evaluated data file includes covariance information for some (not all) of the included evaluations. The subfiles known as the Standards File and Dosimetry File generally provide information on covariances between evaluated cross sections within a set corresponding to a particular reaction. Information on covariances between different reactions is generally not available, although such information is often needed to perform correct analyses. Covariance values are included in the MF = 33 sections. The covariance information is recorded in various formats, and the user must examine the summary documentation (Ref. 11) for the ENDF File rather carefully in order to be able to extract the information properly. ENDF/B-V is the first version of ENDF/B to include some covariance information. As such, there are still serious limitations to the available information, and also some content is of questionable quality. Nevertheless, it represents a start and the reader should have a look into the matter.

It appears that there is a consensus among evaluators to the effect that experimenters should report their experimental errors in some detail in their publications and should make a clear distinction between the uncorrelated errors and the correlated errors. The nature and origin of the correlations should be specified. Some indication as to the sensitivity of the measured results to the experimental parameters would be very useful to evaluators. If ratios are measured, then ratios should be reported in the paper. The evaluators generally agree that it is best if experimenters abstain from evaluating covariance matrices for their own results and leave that job for the evaluators. The most important reason offered is that information may be lost if the experimenter does the job. Evaluators may have difficulty in unravelling the effects of future changes in certain parameters and the errors in these parameters if there are modifications with the passage of time (e.g., decay constants). Another reason is that evaluators wish to retain the option of generating covariance matrices of size suitable for their own evaluation projects. The necessary square size of an experimental covariance matrix rarely needs to equal the number of data points reported by the experimenter. Correlations often involve regions defined by energy limits, or by some other characteristic of the experiment. Each region may involve several data points. Under these conditions, a covariance matrix of square size equal to the total number of data points would contain a number of redundant elements. This viewpoint seems to be quite reasonable, but there may well be some specific cases where it is worthwhile for the experimenter to evaluate his own covariance matrices.

Concerning the use of covariance formalism in analyzing error propagation, one hears questions asked like: Is it necessary? Is the use of covariance analysis a matter of judgment? Since the available covariance information is restricted and is often of dubious accuracy, is application of the method more or less impractical on the basis of the "garbage-in leads to garbage-out" principle? Does it make sense to devote as much time to error analysis as to the performance of the experiment and analysis of the results? Clearly, these are valid questions! Yes, it is a matter of judgment as to how one applies the method. This is true for all endeavors. If there are strong correlations in the data, and these are neglected, then the calculated errors will not be very accurate. If the correlations are weak, neglecting them may be quite acceptable. The matter for judgment is that of deciding just how accurately one needs to know the errors, and, for that matter, just how accurately the uncertainties and correlations for the basic parameters are known. One may well be forced to accept a compromise between accuracy for the results sought and time available to do the job. Applied nuclear data research, like politics, is to a large extent "the art of the possible".

VII. COVARIANCE MATRICES AND THE LEAST-SQUARES METHOD FOR SOLVING OVER-DETERMINED SYSTEMS OF LINEAR EQUATIONS

The reader who is weary, or has run out of time and is looking for a natural stopping point in this treatment of covariance matrices could well stop at the end of the previous section; some new aspects of covariance matrix usage will be considered in this section and those to follow. But, the reader who quits now will surely miss a most interesting and important part of the covariance story. The development and application of proper evaluation techniques is of considerable importance. The data base for applications has expanded greatly during the last decade. It is time to assess these experimental results and make as much use of them as possible before contemplating new measurements of some of these quantities. The situation has been summarized well by J. J. Schmidt (Ref. 12). It will be seen in Section XI that covariance methodology can also be very useful in routine data analysis applications. The reader should invest time to learn this even if no professional evaluation activities are anticipated.

Everybody is familiar with the idea of a system of I linear equations with I unknowns:

$$\begin{aligned} Y_1 &= A_{11}X_1 + A_{12}X_2 + \ldots + A_{1I}X_I \\ &\vdots \\ Y_k &= A_{k1}X_1 + A_{k2}X_2 + \ldots + A_{kI}X_I \\ &\vdots \\ Y_I &= A_{I1}X_1 + A_{I2}X_2 + \ldots + A_{II}X_I \end{aligned} \tag{101}$$

In vector notation, we write

$$\vec{Y} = \bar{A} \bullet \vec{X} \tag{102}$$

It is well known that not all such systems of equations have solutions. Some of the equations in a system may give redundant information so that the system is <u>under-determined</u>. Then, one ends up with a family of solutions. For example,

$$\begin{aligned} 3 &= 2\,X_1 + X_2 \\ 6 &= 4\,X_1 + 2\,X_2 \end{aligned} \tag{103}$$

is an under-determined system of equations with

$$\bar{A} = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \quad . \tag{104}$$

The <u>determinant</u> of $\bar{A}$, designated $|\bar{A}|$ is given by

$$|\bar{A}| = (2)(2) - (1)(4) = 0, \tag{105}$$

so $\bar{A}$ is a singular matrix and the system has no unique solution, but rather a family of solutions given by the equation

$$X_2 = 3 - 2X_1. \tag{106}$$

Consider now the set of equations

$$\begin{aligned} 3 &= 2X_1 + X_2 \\ 4 &= 2X_1 + X_2 \quad . \end{aligned} \tag{107}$$

By inspection, this is <u>clearly</u> an <u>inconsistent</u> set of equations with no <u>exact</u> solution. As before, $|\bar{A}| = 0$, but since there are non-zero elements in $\bar{A}$, $\bar{A}$ has <u>rank</u> > 0 (in fact the rank is 1). But, have a look at the augmented matrix $(\bar{A},\vec{Y})$:

$$(\bar{A},\vec{Y}) = \begin{bmatrix} 2 & 1 & 3 \\ 2 & 1 & 4 \end{bmatrix} \quad . \tag{108}$$

There are 2X2 submatrices of $(\bar{A},\vec{Y})$ which have non-zero determinants, e.g.,

$$\bar{Q} = \begin{bmatrix} 2 & 3 \\ 2 & 4 \end{bmatrix} \quad , \tag{109}$$

with $|\bar{Q}| = 2$. Therefore, the rank of $(\bar{A},\vec{Y})$ is 2. Whenever the rank of the augmented matrix $(\bar{A},\vec{Y})$ differs from the rank of $\bar{A}$, the equations are inconsistent. It is a general theorem that the system represented by Eq. (101) or (102) is consistent if matrix $\bar{A}$ and its augmented matrix $(\bar{A},\vec{Y})$ have the same rank, or, in other words, if the dimensionality of the largest nonsingular submatrices in each are the same. If $\bar{A}$ is nonsingular, then the system automatically is consistent and has a unique solution because the rank of $\bar{A}$ is I and the rank of $(\bar{A},\vec{Y})$ obviously must be I too. the solution is given by

$$\vec{X} = \bar{A}^{-1} \bullet \vec{Y} \quad , \tag{110}$$

where $\bar{A}^{-1}$ is the inverse of $\bar{A}$ which is guaranteed to exist by virtue of the nonsingularity of $\bar{A}$. By definition,

$$\bar{A}^{-1} \bullet \bar{A} = \bar{I} \quad , \tag{111}$$

where $\bar{I}$ is the identity or unit matrix.

In applications, one is often called upon to deduce a <u>best-value</u>, or set of best-values, from available experimental information. This is the definition of <u>evaluation</u>. If there is only one experimental value available in the World for some particular parameter, then it is the best value by default. If there are several values available, generally differing somewhat from each other (within limits) and possessing errors (or, hopefully, more complete covariance information), then the evaluation process is not so obvious but may still be manageable. One possible approach is the generalized linear least-squares method.

The mathematical problem to be addressed is the following: Find the best solution $\vec{X}$ $(X_1, \ldots, X_I)$, in the least-squares sense, to the following overdetermined set of approximate equations which relate the results $\vec{X}$ to a set of available quantities $\vec{Y}(Y_1, \ldots, Y_M)$. Assume that there is covariance matrix $\bar{V}$ which provides the errors, and their correlations, for $\vec{Y}$:

$$\begin{aligned}
Y_1 &\approx A_{11}X_1 + A_{12}X_2 + \ldots + A_{1I}X_I \\
&\vdots \\
Y_k &\approx A_{k1}X_1 + A_{k2}X_2 + \ldots + A_{kI}X_I \\
&\vdots \\
Y_M &\approx A_{M1}X_1 + A_{M2}X_2 + \ldots + A_{MI}X_I \quad .
\end{aligned} \tag{112}$$

In vector notation,

$$\vec{Y} \approx \bar{A} \bullet \vec{X} \tag{113}$$

The symbol "≈" is important because it emphasizes the approximate relation between observables $\vec{Y}$ and the solution $\vec{X}$. The least-squares method says that one way to find the best possible solution to this problem is to obtain the vector $\vec{X}$ which minimizes the chi-square χ^2. The χ^2 for this generalized problem is the quadratic form

$$\chi^2 = (\vec{Y} - \bar{A} \bullet \vec{X})^T \bullet \bar{V}^{-1} \bullet (\vec{Y} - \bar{A} \bullet \vec{X}) \geq 0. \tag{114}$$

χ^2 is a non-negative scaler function of all the available input. One sees immediately that the solution of the problem is going to involve a determination of the inverse of the covariance matrix. Calculation of the inverse of a matrix is a time-consuming operation if the dimension of the matrix is large, so one of the major tasks involved in applying the least-squares method to evaluations is that of finding tricks for keeping the dimensions of the matrices to be inverted as small as possible. One very useful trick, if the correlations are confined to N subgroups of $\vec{Y}$ with no correlations between groups, is to organize the problem so that $\bar{V}$ can be expressed in the form:

$$\bar{V} = \begin{bmatrix} \bar{V}_1 & & & 0 \\ & \ddots & & \\ & & \bar{V}_k & \\ & & & \ddots \\ 0 & & & & \bar{V}_N \end{bmatrix} . \tag{115}$$

Then, the reader can easily verify that

$$\bar{V}^{-1} = \begin{bmatrix} \bar{V}_1^{-1} & & & 0 \\ & \ddots & & \\ & & \bar{V}_k^{-1} & \\ & & & \ddots \\ 0 & & & & \bar{V}_N^{-1} \end{bmatrix} \tag{116}$$

is the desired inverted matrix. The smaller-sized $\bar{V}_k$ matrices can be inverted much more quickly than the full matrix. The covariance matrix will be M × M and is non-singular provided that the available input values $\vec{Y}$ are the result of M distinct observations (i.e., no Y_k is derived from some combination of the

other available values). Thus, the inverse $\bar{V}^{-1}$ exists if the job of setting up the problem was properly done.

How do we find the $\vec{X}$ that minimizes χ^2 in Eq. (114)? It is not too hard to derive the correct answer by an approach which should be sufficiently rigorous to please most Physicists. This proof is presented here in some detail because it gives insight into the basic concept of linear least-squares analysis.

The quantities $\bar{A}$, $\vec{Y}$ and $\bar{V}$ (or $\bar{V}^{-1}$) appearing in Eq. (114) are fixed by the nature of the problem. Only $\vec{X}$ is a variable. Let us write Eq. (114) in the form:

$$(\chi^2)_{\vec{X}} = F(\vec{X}) = (\vec{Y} - \bar{A} \bullet \vec{X})^T \bullet \bar{V}^{-1} \bullet (\vec{Y} - \bar{A} \bullet \vec{X}). \tag{117}$$

what happens to χ^2 if we change $\vec{X}$ by an increment $d\vec{X}$? To answer this, we have to gather courage and plunge into the Calculus (involving matrices). We can write:

$$(\chi^2)_{\vec{X} + d\vec{X}} = F(\vec{X} + d\vec{X}) = [\vec{Y} - \bar{A} \bullet (\vec{X} + d\vec{X})]^T \bullet \bar{V}^{-1} \bullet [\vec{Y} - \bar{A} \bullet (\vec{X} + d\vec{X})] \tag{118}$$

Let's do some matrix algebra:

$$\begin{aligned} F(\vec{X} + d\vec{X}) &= [\vec{Y} - \bar{A} \bullet \vec{X} - \bar{A} \bullet d\vec{X}]^T \bullet \bar{V}^{-1} \bullet [\vec{Y} - \bar{A} \bullet \vec{X} - \bar{A} \bullet d\vec{X}] \\ &= (\vec{Z} - \bar{A} \bullet d\vec{X})^T \bullet \bar{V}^{-1} \bullet (\vec{Z} - \bar{A} \bullet d\vec{X}), \\ &= [\vec{Z}^T - (\bar{A} \bullet d\vec{X})^T] \bullet \bar{V}^{-1} \bullet [\vec{Z} - (\bar{A} \bullet d\vec{X})] \end{aligned} \tag{119}$$

if we define, for convenience,

$$\vec{Z} = \vec{Y} - \bar{A} \bullet \vec{X} \quad . \tag{120}$$

Then, remembering that matrix multiplication is <u>not</u> commutative,

$$F(\vec{X} + d\vec{X}) = \vec{Z}^T \bullet \bar{V}^{-1} \bullet \vec{Z}$$

$$- [\vec{Z}^T \bullet \bar{V}^{-1} \bullet (\bar{A} \bullet d\vec{X}) + (\bar{A} \bullet d\vec{X})^T \bullet \bar{V}^{-1} \bullet \vec{Z}] \tag{121}$$

$$+ (\bar{A} \bullet d\vec{X})^T \bullet \bar{V}^{-1} \bullet (\bar{A} \bullet d\vec{X}) \quad .$$

The third line on the right-hand side of Eq. (121) is second-order in $d\vec{X}$ and will have a negligible magnitude as $d\vec{X}$ approaches zero, relative to the other terms. Therefore, in the tradition of the Calculus, we can neglect it. The first line on the right-hand side of Eq. (121) is seen from Eqs. (117) and (120) to be simply $(\chi^2)_{\vec{X}}$ or $F(\vec{X})$. Thus, Eq. (121) becomes:

$$d(\chi^2)_{\vec{X}} = dF(\vec{X}) = F(\vec{X} + d\vec{X}) - F(\vec{X})$$

$$= - \vec{Z}^T \bullet \bar{V}^{-1} \bullet (\bar{A} \bullet d\vec{X}) - (\bar{A} \bullet d\vec{X})^T \bullet \bar{V}^{-1} \bullet \vec{Z}) \tag{122}$$

$$= - [(\vec{Z}^T \bullet \bar{V}^{-1}) \bullet (\bar{A} \bullet d\vec{X})] - [(\vec{Z}^T \bullet \bar{V}^{-1})^T \bullet (\bar{A} \bullet d\vec{X})]^T \ .$$

We used two laws of matrix algebra to arrive at this result:

One rule states that for any two arbitrary matrices $\bar{C}_1$ and $\bar{C}_2$,

$$(\bar{C}_1 \bullet \bar{C}_2)^T = \bar{C}_2{}^T \bullet \bar{C}_1{}^T \quad . \tag{123}$$

The second rule states that for our nonsingular symmetric covariance matrix,

$$(\bar{V}^{-1})^T = (\bar{V}^T)^{-1} = \bar{V}^{-1}, \tag{124}$$

so the inverse $\bar{V}^{-1}$ of $\bar{V}$ is also symmetric.

But, the transpose of a scalar quantity is just the scaler quantity itself, and Eq. (122) is a scalar equation. Therefore, Eq. (122) becomes

$$d(\chi^2)_{\vec{X}} = dF(\vec{X}) = - 2(\vec{Z}^T \bullet \bar{V}^{-1} \bullet \bar{A}) \bullet d\vec{X} \tag{125}$$

The quantity χ^2 has a minimum when its differential is zero for small variations $d\vec{X}$ about the minimum. Therefore,

$$\vec{Z}^T \bullet \bar{V}^{-1} \bullet \bar{A} = 0 \tag{126}$$

is the condition for a minimum in χ^2.

Applying the rules of matrix algebra given above, we get

$$\begin{aligned} \vec{Z}^T \bullet \bar{V}^{-1} \bullet \bar{A} &= \vec{Z}^T \bullet (\bar{V}^{-1} \bullet \bar{A}) \\ &= [(\bar{V}^{-1} \bullet \bar{A})^T \bullet \vec{Z}]^T = (\bar{V}^{-1} \bullet \bar{A})^T \bullet \vec{Z} \\ &= \bar{A}^T \bullet \bar{V}^{-1} \bullet \vec{Z} = 0. \end{aligned} \tag{127}$$

Substitution of Eq. (120) into Eq. (126) leads to

$$\bar{A}^T \bullet \bar{V}^{-1} \bullet \vec{Y} = (\bar{A}^T \bullet \bar{V}^{-1} \bullet \bar{A}) \bullet \vec{X} \tag{128}$$

The matrix $(\bar{A}^T \bullet \bar{V}^{-1} \bullet \bar{A})$ is nonsingular with size I × I, therefore, there exists an inverse which we can call $\bar{C}$:

$$\bar{C} = (\bar{A}^T \bullet \bar{V}^{-1} \bullet \bar{A})^{-1}. \tag{129}$$

The desired least-squares solution for $\vec{X}$ is thus:

$$\vec{X} = \bar{C} \bullet \bar{A}^T \bullet \bar{V}^{-1} \bullet \vec{Y} \tag{130}$$

The matrix $\bar{C}$ is the covariance matrix for the solution $\vec{X}$. Since I < M, the calculation of $\bar{C}$ is generally not as big a problem as evaluation of $\bar{V}^{-1}$.

Important information can be deduced from the minimal χ^2, calculated with Eq. (114) using the solution vector $\vec{X}$. χ^2 should be distributed according chi-square tables applicable to the actual number of degrees of freedom f for the problem. Chi-square tables for Normal distributions are available from many sources (e.g., Ref. 9, p. 313). As an example, assume that 25 items of data (M) are evaluated by the least-squares method in order to yield a result consisting of 15 best values (I), with a χ^2 of 7.92 according to Eq. (114). The number of degrees of freedom f (= M-I) in this problem is 10. According to Table C-4, p. 313 of Ref. 9, the probability of observing a χ^2 this size for the evaluation problem is in the range 60-70% which is reasonable. A crude interpretation of this result

is that the scatter of the input values is consistent with the assigned errors. A large resultant χ^2 with low probability would imply an inconsistency between the actual scatter of input values and the errors claimed for them (perhaps the errors claimed were too small).

What should be done if a large χ^2 with low probability is obtained from a least-squares analysis. The matter has been touched upon by Peele (Ref. 3), Mannhart (Ref. 7) and Poenitz (Ref. 8). If there are several degrees of freedom f, it is reasonable to expect approximately 50% probability for χ^2/f to be near unity. If $\chi^2/f \gg 1$, then one possible approach is to make an adjustment to the solution covariance matrix $\overline{C}$ by multiplying it by χ^2/f. This is equivalent to multiplying the covariance matrix $\overline{V}$ for the input data by the same factor. There may be justification for this step if the evaluator feels that the input errors were underestimated, but it is a crude "adjustment" at best. If χ^2/f is much less than unity, one should not reduce the errors. Problems caused by shape inconsistencies rather than excess scatter should not be dealt with by scaling the covariance matrices.

Example: Evaluation of a single cross section

An evaluator wishes to obtain the best value for a cross section σ, and its error. He scans the literature and finds two values with their reported errors as follows: $S_1 = 1.85$ b(±6%) and $S_2 = 1.94$ b(±8%). Based on the available information, he estimates that the errors are approximately 50% correlated.

Since the evaluator seeks a cross section, and is given cross section data as input, the equivalent of Eq. (112), with $\vec{Y}$ replaced by vector $\vec{S}$ consisting of S_1 and S_2, and $\vec{X}$ of σ only, is:

$$S_1 \approx \sigma$$
$$S_2 \approx \sigma \,, \tag{131}$$

with the $\overline{A}$ matrix quite simply

$$\overline{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} . \tag{132}$$

The covariance matrix $\overline{V}$ is:

$$\overline{V} = 10^{-2} \begin{bmatrix} 1.232 & 0.8614 \\ 0.8614 & 2.409 \end{bmatrix} . \tag{133}$$

The corresponding inverse matrix is:

$$\overline{V}^{-1} = 10^2 \begin{bmatrix} 1.082 & -0.03870 \\ -0.3870 & 0.5535 \end{bmatrix} . \tag{134}$$

Therefore,

$$(\overline{A}^T \bullet \overline{V}^{-1} \bullet \overline{A}) = 10^2[1,1] \bullet \begin{bmatrix} 1.082 & -0.3870 \\ -0.3870 & 0.5535 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{135}$$

$$= \overline{C}^{-1} = 86.15,$$

a simple scalar. The covariance "matrix" for the evaluated cross section is the reciprocal (inverse) of this scalar. Therefore

$$\overline{C} = C = 1.161 \times 10^{-2} \tag{136}$$

The solution σ (evaluated cross section) is given by

$$\sigma = C\ \overline{A}^T \bullet \overline{V}^{-1} \bullet \vec{S}$$

$$= (1.161 \times 10^{-2})\ (10^2)\ [1,1] \bullet \begin{bmatrix} 1.082 & -0.3870 \\ -0.3870 & 0.5535 \end{bmatrix} \bullet \begin{bmatrix} 1.85 \\ 1.94 \end{bmatrix} \tag{137}$$

$$= 1.868\text{b} \quad .$$

The error in σ is given by $C^{1/2} \approx 0.1077$b. So, the evaluation procedure has given as the best estimated value, the result $\sigma = 1.868$ b(±5.8%). This solution leads to a χ^2 of 0.3220 which indicates reasonable consistency (50-60% likelihood range) for the input data.

If we had ignored the correlations in the errors of the input data, and performed an evaluation using the formalism of this section, we would have obtained the value $\sigma = 1.880$b(±4.8%), which is just the weighted average of the two results as derived from the expression

$$\bar{\sigma} \approx \frac{(1.85/6^2) + (1.94/8^2)}{(1/6^2) + (1/8^2)} \quad . \tag{138}$$

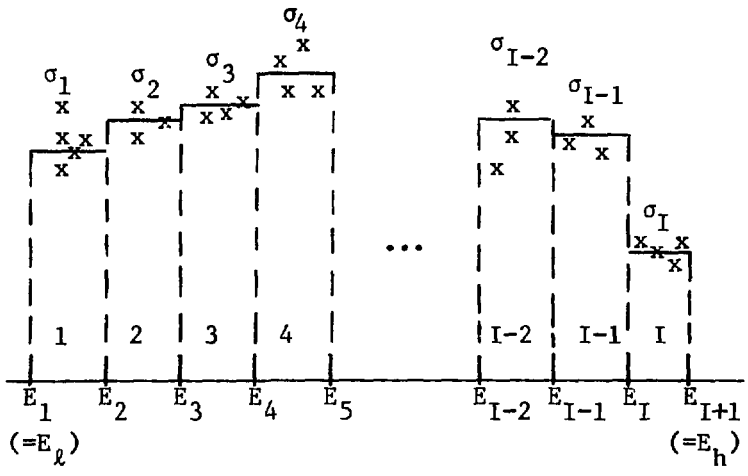
The use of covariances in the generalized linear least-squares method leads to a better estimation of the cross section. The formalism also provides the covariance (error) for the evaluated result, and a χ^2 which measures the consistency of the input data.

The error which is obtained when the correlation element is ignored is smaller in this example than the error obtained when the correlation element is included. It has been claimed that the inclusion of complete covariance information in evaluations reduces the error. Apparently this is not always a true statement. It does not prove to be the case in the example here.

VIII. ONE METHOD FOR GENERATING EVALUATED CROSS SECTIONS FROM A DATA BASE USING THE LEAST-SQUARES METHOD

The basic rules of least-square evaluation are straightforward, but realistic problems are usually sufficiently complex to require approximations and assumptions to be made to reduce the labor to manageable levels. Special techniques need to be developed for most specific applications. This section will deal with a particular type of problem which is more realistic than the simple example of the previous section, and a least-squares method for solving it will be indicated.

Suppose we want to represent a cross section excitation function by I group average values $\vec{\sigma}$ over an energy range (E_ℓ, E_h). Energy limits for the groups are selected on the basis of available experimental data for the groups.



No apriori knowledge of the true cross section (shape or normalization) is available, so we must do the best we can with the experimental data on hand. Assume that there are J different experimental cross section sets available from the literature for the range (E_ℓ, E_h). Assume that the data base and selection of group energy limits $(E_1, \ldots, E_{I+1})$ are such that no group is void of experimental data. No information about the structure of the cross section within a specific group is sought. Hence, each measured value is considered to be an approximation to the group average value for the group where it lies. If there is only one experimental value available for a particular group, then the derived group value may be very uncertain. However if there are many experimental points scattered in energy throughout the group, this least-squares method should give a fairly reliable group-average value. The reader who is familiar with Monte Carlo techniques will recall that integrals can be calculated by sampling functional values within the specified limits of integration. The group average cross section is derived from the integral of the true excitation function over the group interval. The experimental values correspond to a sampling of the cross section at various energies within the group interval.

The evaluation effort is considerably reduced for this problem if the correlations between the various data sets are small enough to neglect. This may not be true for all the data sets involved, but surely it is very likely that it is true for some of them. By judicious organization of the input values, one tries to create a covariance matrix $\bar{V}$ for the input data which has large blocks of zero off-diagonal matrices. Generally those uncertainties associated with specific aspects of individual experiments are uncorrelated with those for other experiments, and these uncertainties often are dominant contributors to the total error. Common or strongly correlated standards can introduce sizable correlations for data sets unless the standard error is small (e.g. the hydrogen cross section).

Before beginning the evaluation, the input data should be cleaned up by performing the following operations as required:

i) Correct all data to latest values of decay constants, standard cross sections, etc.

ii) If the errors for two cross section sets are so strongly correlated that neglect of the correlation would be significant, then consider these two sets together.

iii) Construct the covariance matrices for each set (if not available), and for coupled sets as per (ii) above.

etc...

An example of the data preparation process appears in a recent report by Tagesen et al. (Ref. 13). Poenitz (Ref. 8) also discusses this topic. Strictly speaking, it is not correct to evaluate cross sections for an individual reaction alone as implied in this example. However, it is not always practical to perform simultaneous evaluations for several reactions.

The evaluation proper begins when the evaluator collects all the prepared data into K sets ($K \leq J \leq M$) where M is the total number of available cross section points. The K sets are represented by the vector

$$\vec{S} = \begin{bmatrix} \vec{S}_1 \\ \vdots \\ \vec{S}_K \end{bmatrix}, \tag{139}$$

with covariance matrix $\bar{V}$ of the form

$$\bar{V} = \begin{bmatrix} \bar{V}_1 & & 0 \\ & \ddots & \\ 0 & & \bar{V}_K \end{bmatrix}. \tag{140}$$

The K sets are formed so that the errors for each set are considered uncorrelated to those for the other sets. The dimensionalities (number of experimental values in each set) are $M_1, \ldots, M_K$ with sum equal to M. We could consider keeping M as small as possible by averaging all the data from a given experiment which falls in the same group to one value. From Eqs. (115) and (116),

$$\bar{V}^{-1} = \begin{bmatrix} \bar{V}_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \bar{V}_K^{-1} \end{bmatrix} . \tag{141}$$

An MxI matrix $\bar{A}$ can be expressed in terms of K submatrices $\bar{A}_k$ with dimension M_kxI as follows:

$$\bar{A} = \begin{bmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_K \end{bmatrix} , \tag{142}$$

The elements of the $\bar{A}_k$ matrices are either zero or unity depending on how the experimental data relate to a particular group. For example, the element $(\bar{A}_k)_{mk,i}$ is unity if the m_k-th value for experimental set k corresponds to group i, while $(\bar{A}_k)_{mk,i'}$ is zero for all $i' \neq i$.

Then, according to the formalism of the preceding section, the solution vector $\vec{\sigma}$ (I evaluated group cross sections) is given by:

$$\vec{\sigma} = \bar{C} \bullet \bar{A}^T \bullet \bar{V}^{-1} \bullet \vec{S}$$

$$= \bar{C} \bullet (\bar{A}_1^T \bullet \bar{V}_1^{-1} \bullet \vec{S}_1 + \ldots + \bar{A}_K^T \bullet \bar{V}_K^{-1} \bullet \vec{S}_K) \quad , \tag{143}$$

while the covariance matrix $\bar{C}$ for the evaluation is deduced from

$$\bar{C}^{-1} = \bar{A}^T \bullet \bar{V}^{-1} \bullet \bar{A}$$

$$= \bar{A}_1^T \bullet \bar{V}_1^{-1} \bullet \bar{A}_1 + \ldots + \bar{A}_K^T \bullet \bar{V}_K^{-1} \bullet \bar{A}_K . \tag{144}$$

Derivation of Eqs. (143) and (144) is a straightforward problem of matrix manipulation which originates from the rules supplied by Eqs. (129) and (130). So, the evaluation problem-exclusive of data preparation-consists of inverting K matrices $\overline{V}_1, \ldots, \overline{V}_K$ with squares sizes $M_1, \ldots, M_K$, and one matrix $\overline{C}^{-1}$ with square sizes I.

IX. USE OF AN APRIORI AND LINEARIZATION OF NON-LINEAR LEAST-SQUARES PROBLEMS

There is an inherent limitation to the approach taken in the last two sections. It requires a linear relationship between the input observables, $\vec{y}$, and the solution to the problem, $\vec{x}$, as expressed, for example, by Eq. (112). This excludes, quite clearly, the possibility for considering observed ratios of quantities in the evaluation procedure. Since much useful nuclear data comes to us in the form of ratios and other nonlinear forms, this is not a tolerable limitation. In this section, we look at another approach to the least-square problem which enables us to bypass this limitation. This involves use of what is known as the apriori estimate. It is an approximation method.

Suppose that one has available a set of measured quantities $\vec{\xi}$ (ξ_1, ..., ξ_M) with a covariance matrix $\bar{V}$. Assume that each of these measured quantities approximates a function f_m of the desired evaluated cross section set $\vec{\sigma}$ (σ_1, ..., σ_I) according to

$$\xi_m \approx f_m(\vec{\sigma}) \quad (m = 1, M) \quad . \tag{145}$$

Now, suppose we have an apriori estimate of $\vec{\sigma}$ which we designate $\vec{\sigma}_0$ (σ_{10}, ..., σ_{10}). Each value $f_m(\vec{\sigma})$ can be related to the corresponding $f_m(\vec{\sigma}_0)$ by means of the approximation

$$f_m(\vec{\sigma}) \approx f_m(\vec{\sigma}_0) + \sum_{i=1}^{I} \left(\frac{\partial f_m}{\partial \sigma_i}\right)_0 (\sigma_i - \sigma_{i0}) \quad , \tag{146}$$

which is a truncated Taylor's series expansion. Next, define vectors $\vec{x}$ and $\vec{y}$ by means of

$$x_i = \sigma_i - \sigma_{0i} \qquad (i = 1, I) \quad , \tag{147}$$

$$y_m = \xi_m - f_m(\vec{\sigma}_0) \qquad (m = 1, M) \quad . \tag{148}$$

Then Eq. (145) can be expressed in the form

$$y_m \approx \sum_{i=1}^{I} \left(\frac{\partial f_m}{\partial \sigma_i}\right)_0 x_i \qquad (m = 1, M) \quad . \tag{149}$$

Equation (149) is identical to Eqs. (101) or (102) provided that the elements of the matrix $\bar{A}$ are defined as the partial derivatives

$$A_{mi} = \left(\frac{\partial f_m}{\partial \sigma_i}\right)_0 \qquad (m = 1, M;\ i = 1, I) \quad , \tag{150}$$

calculated using the apriori estimate $\vec{\sigma}_0$. Furthermore, since no error is assumed for the apriori estimate $\vec{\sigma}_0$ in this derivation, the covariance matrix for $\vec{y}$ is also $\bar{V}$ while the covariance matrix $\bar{C}$ for the solution $\vec{x}$ is also the covariance matrix for $\vec{\sigma}$ as derived using Eq. (147). The matrix $\bar{C}$ is calculated using Eq. (129), and the solution vector $\vec{x}$ (and thus $\vec{\sigma}$) is obtained from Eq. (130). Equation (114) can be used to calculate the apriori chi-square and the chi-square provided by the solution. The apriori chi-square χ_0^2 is obtained when $\vec{x}$ is set equal to the zero vector in Eq. (114). The solution χ^2 is obtained when the solution vector $\vec{x}$ from Eq. (130) and the matrix $\bar{A}$ from Eq. (150) are used in Eq. (114).

An improved solution can be sought by means of iteration, and convergence of χ^2 can be used as a test. When the apriori estimate $\vec{\sigma}_0$ is reasonable, one can expect the iterative process to lead to convergence rather quickly (two to three passes at most). Owing to the time consuming nature of matrix inversion, selection of a physically reasonable apriori solution $\vec{\sigma}_0$ is quite important for practical reasons.

If ξ_m approximates a particular cross section σ_i, then

$$f_m(\vec{\sigma}) \approx \sigma_i \quad , \tag{151}$$

$$A_{mi'} = \begin{cases} 1 & \text{if} \quad i' = i \quad , \\ 0 & \text{if} \quad i' \neq i \quad . \end{cases} \tag{152}$$

When ξ_m approximates a ratio of two cross sections in the set, then

$$f_m(\vec{\sigma}) \approx \sigma_i/\sigma_i'. \tag{153}$$

$$A_{mi''} = \begin{cases} 1/\sigma_i' & \text{if} \quad i'' = i \quad , \\ -\sigma_i/\sigma_{i'}^2 & \text{if} \quad i'' = i' \quad , \\ 0 & \text{if} \quad i'' \neq i \text{ or } i' \quad . \end{cases} \tag{154}$$

Using this formalism, ratios are as easy to handle as cross sections.

X. HOW TO UPDATE AN EVALUATION WITH NEW DATA WITHOUT PERFORMING A COMPLETELY NEW EVALUATION

Often an evaluator will be faced with the task of updating an existing evaluation by including one or more new data sets. He would like to not have to redo the entire evaluation. This section addresses this special problem.

First, we need to define the term relative covariance matrix. Suppose $\bar{M}_x$ is the actual covariance matrix applying to a set of quantities $\vec{x}$. Then, the relative covariance matrix $\bar{M}$ corresponding to $\bar{M}_x$ is deduced from the formula

$$M_{xij} = x_i x_j M_{ij} \quad \text{(all i and j)}. \tag{155}$$

The elements of $\bar{M}$ are therefore dimensionless.

Suppose we have an evaluated group cross section set $\vec{\sigma}_0(\sigma_{01},\ldots, \sigma_{0I})$ with a corresponding relative covariance matrix $\bar{C}_0$. A set of new experimental values $\vec{\xi}(\xi_1\ldots,\xi_2)$ with a relative covariance matrix $\bar{V}_\xi$ becomes available. We wish to generate a new evaluation $\vec{\sigma}_1(\sigma_{11},\ldots, \sigma_{1I})$ with a corresponding relative covariance matrix $\bar{C}_1$. How can this be done?

Consider the I cross sections σ_0 as input data with relative covariance matrix $\bar{C}_0$ and the L new values as additional input data with relative covariance matrix $\bar{V}_\xi$. Also, consider the apriori function $\sigma_0(E)$ defined by $\vec{\sigma}_0$ itself. Then, define a new vector $\vec{y}$ (with dimension I + L) by

$$y_m = \left(\frac{\sigma_{0i} - \sigma_{0i}}{\sigma_{0i}}\right) = 0, \ (m = 1,\ldots, I),$$

$$y_{I+\ell} = \left(\frac{\xi_\ell - \xi_{\ell 0}}{\xi_{\ell 0}}\right)(\ell = 1,\ldots, L). \tag{156}$$

For this method to work we must assume that the new data set is uncorrelated to the original evaluation. Then, the actual covariance matrix $\bar{V}_y$ for $\vec{y}$ is simply

$$\bar{V} = \begin{bmatrix} \bar{C}_0 & \bar{0} \\ \bar{0} & \bar{V}_\xi \end{bmatrix}. \tag{157}$$

We seek $\vec{x}_1$ defined by

$$x_{1i} = \left(\frac{\sigma_{1i} - \sigma_{0i}}{\sigma_{0i}}\right)(i = 1,\ldots, I). \tag{158}$$

The solution to this problem is derived from

$$\vec{x}_1 = \bar{C}_1 \bullet \bar{A}^T \bullet \bar{V}_y^{-1} \bullet \vec{y}, \tag{159}$$

$$\bar{C}_1^{-1} = \bar{A}^T \bullet \bar{V}_y^{-1} \bullet \bar{A}, \tag{160}$$

where

$$\bar{A} = \begin{bmatrix} \bar{A}_0 \\ \\ \bar{A}_\xi \end{bmatrix} \tag{161}$$

$\bar{A}_0$ is simply an I × I dimensional unit matrix as can be seen from an inspection of Eqs. (156) and (158). Since we linearize the problem by using the previous evaluation as an apriori function, $\bar{A}_\xi$ will be a matrix consisting of elements whose values depend upon the character of the new information.

Equation (159) can be written in the form:

$$\vec{x}_1 = \bar{C}_1 \bullet \begin{bmatrix} \bar{A}_0 \\ \\ \bar{A}_\xi \end{bmatrix}^T \bullet \begin{bmatrix} \bar{C}_0^{-1} & \bar{0} \\ \\ \bar{0} & \bar{V}_\xi^{-1} \end{bmatrix} \bullet \begin{bmatrix} \vec{0} \\ \\ \vec{y}_\xi \end{bmatrix}$$

$$= \bar{C}_1 \bullet \bar{A}_\xi^T \bullet \bar{V}_\xi^{-1} \bullet \vec{y}_\xi \tag{162}$$

where $\vec{y}_\xi$ is the subvector of $\vec{y}$ with elements $y_{I+1},\ldots, y_{I+L}$. The rest of $\vec{y}$ consists of zero elements according to Eq. (156). The matrix $\bar{C}_1$ is derived from the relation

$$\bar{C}_1^{-1} = \begin{bmatrix} \bar{A}_0 \\ \bar{A}_\xi \end{bmatrix}^T \bullet \begin{bmatrix} \bar{C}_0^{-1} & \bar{0} \\ \bar{0} & \bar{V}_\xi^{-1} \end{bmatrix} \bullet \begin{bmatrix} \bar{A}_0 \\ \bar{A}_\xi \end{bmatrix} \tag{163}$$

Equation (163) leads to the result

$$\bar{C}_1^{-1} = \bar{A}_0^T \bullet \bar{C}_0^{-1} \bullet \bar{A}_0 + \bar{A}_\xi^T \bullet \bar{V}_\xi^{-1} \bullet \bar{A}_\xi$$

$$= \bar{C}_0^{-1} + \bar{A}_\xi^T \bullet \bar{V}_\xi^{-1} \bullet \bar{A}_\xi \quad , \tag{164}$$

since $\bar{A}_0$ is and I × I unit matrix. Clearly, calculations based on Eqs. (162) and (164) involve much less work than would be encountered if the entire evaluation were repeated. The vector $\vec{\sigma}_1$ is derived from

$$\sigma_{1i} = \sigma_{0i}(1 + x_{1i}) \quad (i = 1,\ldots, I) \quad , \tag{165}$$

and the matrix $\bar{C}_1$ is the relative covariance matrix for the updated evaluation.

The method indicated here can be generalized to include several new, independent data sets. The reader should be able to convince himself that J new data sets lead to a solution of the form:

$$\vec{x}_1 = \bar{C}_1 \bullet \left(\sum_{j=1}^{J} \bar{A}_{\xi j}^T \bullet \bar{V}_{\xi j}^{-1} \bullet \vec{y}_{\xi j} \right),$$

$$\bar{C}_1^{-1} = \bar{C}_0^{-1} + \sum_{j=1}^{J} \bar{A}_{\xi j}^T \bullet \bar{V}_{\xi j}^{-1} \bullet \bar{A}_{\xi j} \quad , \tag{166}$$

where the subscripts j apply to the data sets j = 1,..., J.

Although we have considered both linear and nonlinear least-squares problems in Sections VII-X, this discussion has not exhausted the possibilities of the least-squares method. More general formulations can be developed which allow, e.g., even the adjustment of elements in $\vec{y}$ and in the <u>transfer matrix</u> $\bar{A}$ from Eq. (113) in order to satisfy a least squares condition. Perey (Ref. 4) has discussed such formulations in his treatment of spectrum unfolding for dosimetry applications. This subject will not be pursued here, so the reader should refer to Ref. 4 if he is interested in some of these other possibilities.

XI. USE OF THE LEAST-SQUARES METHOD IN DATA FITTING

Sections VIII through X emphasize applications of the least-squares method to data evaluation. Formal data evaluation is very probably an infrequent activity for most experimentalists, but there are many applications of the method described in Section VII to everyday data processing activities. The objective of this section is to indicate the potential of the method and to stimulate the reader to pursue the matter further.

Data averaging and curve fitting are routine activities for experimental Physicists. The least-squares method and covariances can be used to formulate these operations in a very general and flexible way. The technique of linearization by means of an apriori can be used to deal with a variety of non-linear problems as well. The best way to approach this topic is through examples. Four examples will be presented in this section.

Example: Averaging Uncorrelated Data

Suppose that two independent measurements have been made of the same physical quantity yielding (y_1, E_{y1}) and (y_2, E_{y2}) for the experimental values and their errors. The independence of the measurements insures that the errors are uncorrelated. What is the best estimate (x, E_x) for this physical quantity and its error based on the available data?

The method of least squares is directly applicable to this problem. It is a rather elementary evaluation exercise.

The equivalent of Eq. (112) is

$$y_1 \approx x$$

$$y_2 \approx x. \tag{167}$$

The matrix $\bar{A}$ of coefficients is thus quite simply

$$\bar{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{168}$$

Likewise, the value vector $\vec{y}$ and covariance matrix $\bar{V}$ are

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \tag{169}$$

$$\bar{V} = \begin{bmatrix} E_{y1}^{2} & 0 \\ 0 & E_{y2}^{2} \end{bmatrix} . \tag{170}$$

$\bar{V}$ is a diagonal matrix since the errors for $\vec{y}$ are uncorrelated. The inverse $\bar{V}^{-1}$ is given by

$$\bar{V}^{-1} = \begin{bmatrix} E_{y1}^{-2} & 0 \\ 0 & E_{y2}^{-2} \end{bmatrix} . \tag{171}$$

We now have all information needed to apply the generalized least-squares method. The required formulas are Eqs. (129) and (130). Here is the explicit analysis:

$$\bar{V}^{-1} \bullet \bar{A} = \begin{bmatrix} E_{y1}^{-2} & 0 \\ 0 & E_{y2}^{-2} \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} E_{y1}^{-2} \\ E_{y2}^{-2} \end{bmatrix} , \tag{172}$$

$$\begin{aligned} \bar{C}^{-1} &= \bar{A}^{T} \bullet \bar{V}^{-1} \bullet \bar{A} = [1,1] \bullet \begin{bmatrix} E_{y1}^{-2} \\ E_{y2}^{-2} \end{bmatrix} \\ &= E_{y1}^{-2} + E_{y2}^{-2} \qquad \text{(a scalar).} \end{aligned} \tag{173}$$

The inverse of the scalar $\bar{C}^{-1}$ is its reciprocal

$$\bar{C} = (E_{y1}^{-2} + E_{y2}^{-2})^{-1} . \tag{174}$$

According to the theory, the error E_x in the derived "best" value x can be extracted from the covariance matrix $\bar{C}$ via

$$E_x^2 = \bar{C} . \tag{175}$$

We therefore obtain the well-known expression

$$E_x^{-2} = E_{y1}^{-2} + E_{y2}^{-2} \tag{176}$$

relating the errors in the measured values to the error in the derived "best" value.

Proceeding with the analysis, we have:

$$\bar{V}^{-1} \bullet \vec{y} = \begin{bmatrix} E_{y1}^{-2} & 0 \\ 0 & E_{y2}^{-2} \end{bmatrix} \bullet \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 \; E_{y1}^{-2} \\ y_2 \; E_{y2}^{-2} \end{bmatrix} . \tag{177}$$

$$\bar{A}^T \bullet \bar{V}^{-1} \bullet \vec{y} = y_1 \; E_{y1}^{-2} + y_2 \; E_{y2}^{-2} , \tag{178}$$

$$x = \vec{x} = \bar{C} \bullet \bar{A}^T \bullet \bar{V}^{-1} \bullet \vec{y} = \left(\frac{y_1 \; E_{y1}^{-2} + y_2 \; E_{y2}^{-2}}{E_{y1}^{-2} + E_{y2}^{-2}}\right). \tag{179}$$

Eq. (179) indicates that the "best" value x is the weighted average. The χ^2 for this averaging process can be derived by inserting the solution into Eq. (114). So, we see that the least squares method enables us to derive a familiar result in a very general way, and it gives a χ^2 value to test the consistency of the input data as well.

Example: Averaging Correlated Data

Refer to the preceding example and assume that the errors E_{y_1} and E_{y_2} are correlated. Such a correlation is actually very likely if y_1, and y_2 have been measured in the same laboratory using similar techniques. Let $\bar{Q}$ be the correlation matrix given by

$$\bar{Q} = \begin{bmatrix} 1 & q_{12} \\ q_{21} & 1 \end{bmatrix} = \begin{bmatrix} 1 & q_{12} \\ q_{12} & 1 \end{bmatrix} , \tag{180}$$

since q_{21} and q_{12} must be equal by the symmetric nature of $\bar{Q}$. Thus, the covariance matrix $\bar{V}$ is

$$\bar{V} = \begin{bmatrix} E_{y1}{}^2 & q_{12}E_{y1}E_{y2} \\ q_{12}E_{y1}E_{y2} & E_{y2}^2 \end{bmatrix} . \tag{181}$$

The analysis needed to derive (x, E_x) proceeds much as it did for the preceding example. The results are:

$$\bar{V}^{-1} = |\bar{V}|^{-1} \begin{bmatrix} E_{y2}^2 & -q_{12}E_{y1}E_{y2} \\ -q_{12}E_{Y1}E_{Y2} & E_{Y1}^2 \end{bmatrix} , \tag{182}$$

with the determinant $|\bar{V}|$ equal to

$$|\bar{V}| = E_{y1}{}^2 \; E_{y2}{}^2 \; (1 - q_{12}{}^2) \quad ,$$

and

$$E_x{}^{-2} = \bar{C}^{-1} = \bar{A}^T \bullet \bar{V}^{-1} \bullet \bar{A} \quad \text{(a scalar)}$$

$$= (1 - q_{12}{}^2)^{-1}(E_{y1}{}^{-2} + E_{y2}{}^{-2} - 2q_{12} \; E_{y1}{}^{-1} \; E_{y2}{}^{-1}), \tag{183}$$

and

$$x = \vec{x} = \bar{C} \bullet \bar{A}^T \bullet \bar{V}^{-1} \bullet \vec{y}$$

$$= \frac{[y_1(E_{y1}{}^{-2} - q_{12}E_{y1}{}^{-1}E_{y2}{}^{-1}) + y_2(E_{y2}{}^{-2} - q_{12}E_{y1}{}^{-1} \; E_{y2}{}^{-1})]}{[(E_{y1}{}^{-2} - q_{12}E_{y1}{}^{-1} \; E_{y2}{}^{-1}) + (E_{y2}{}^{-2} - q_{12}E_{y1}{}^{-1} \; E_{y2}{}^{-2})]} \tag{184}$$

This rather complicated expression could not be deduced very easily by simple arguments. The solution reduces to the result from the preceeding example when q_{12} goes to zero (no correlation). The case where q_{12} approaches unity (fully correlated errors) presents special problems. Matrix $\bar{V}$ approaches singularity so the problem would appear to have no solution. This is not the case as can be seen from a closer examination of the mathematics. The errors E_{y1} and E_{y2} should be written in the form

$$E_{y1} = (e_1{}^2 + E^2)^{1/2}$$
$$E_{y2} = (e_2{}^2 + E^2)^{1/2} \tag{185}$$

where E is the common correlated error and e_1 and e_2 are the statistical errors. Then,

$$q_{12} = E^2/E_{y1}\ E_{y2} \quad . \tag{186}$$

The following limits are approached:

$$q_{12} \xrightarrow[e_1,\ e_2 \ll E]{} 1\ , \tag{187}$$

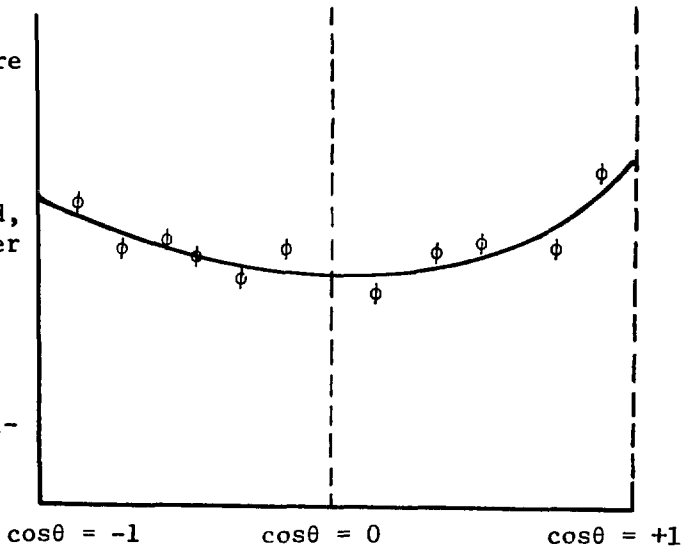
$$E_{y1},\ E_{y2},\ E_x \xrightarrow[e_1,\ e_2 \ll E]{} E, \tag{188}$$

$$x \xrightarrow[e_1,\ e_2 \ll E]{} \frac{y_1\ e_1{}^{-2} + y_2\ e_2{}^{-2}}{e_1{}^{-2} + e_2{}^{-2}} \quad . \tag{189}$$

The physical interpretation of these results is clear. When the statistical error becomes very small, the error in the "best" value approaches the common systematic error. The best value is the weighted average of the measured values, using only the residual statistical errors to calculate weighting factors. When the statistical error becomes totally negligible, $x \approx y_1 \approx y_2$ and Eq. (189) is not needed. Whenever the systematic errors are dominant, plural measurements of a physical quantity will not lead to an improvement in accuracy. Researchers tend to overlook this important fact.

Example: Legendre Expansion Fit to an Angular Distribution

Angular distribution data are usually fitted using Legendre polynomial expansions. Conventional methods generally ignore the possibility of correlations in fitting the data. Using the generalized least-squares method, we can perform the fitting either with or without assumed input correlations. The method will be indicated in this example which considers the fitting of an m-term expansion fit to n data points (n > m). The experimental data points $y_1,\ldots,y_n$ $(\vec{y})$ are assumed to have errors $E_{y1},\ldots,E_{yn}$ $(\vec{E}_y)$ whose correlations are defined by the nxn matrix $\bar{Q}$,



$$\bar{Q} = \begin{bmatrix} 1 & q_{12} & \cdots & q_{1n} \\ q_{21} & 1 & \cdots & q_{2n} \\ \vdots & \vdots & & \vdots \\ q_{n1} & q_{n2} & \cdots & 1 \end{bmatrix}, \tag{190}$$

with

$$q_{ii} = 1, \qquad (i,j = 1,n) \qquad q_{ij} = q_{ji}. \tag{191}$$

If all q_{ij} are zero for $i \neq j$, then we have the case where the errors in the distribution points are assumed to be uncorrelated. The function to be fitted to the data is

$$Y(\theta) = \sum_{j=1}^{m} p_j \, P_{j-1}(\cos\,\theta) \tag{192}$$

where P_{j-1} (cos θ) is the Legendre polynomial of order j - 1. The least-squares problem consists of finding the best set of values $p_1,\ldots,p_m$ ($\vec{p}$) to satisfy the set of approximate equations

$$y_i \approx \sum_{j=1}^{m} p_j \, P_{j-1} \, (\cos \theta_i) \qquad (i = 1,n). \tag{193}$$

It is assumed that the angles θ_i (i=1,n) are precisely known. Eq. (193) is equivalent to Eq. (112), or to Eq. (113) if we define $\bar{A}$ as the nxm matrix

$$\bar{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}, \tag{194}$$

with elements

$$a_{ij} = P_{j-1}(\cos \theta_i) \qquad (i = 1,n \text{ and } j = 1,m), \tag{195}$$

and $\vec{p}$ equivalent to the least-squares solution vector $\vec{x}$. The covariance matrix $\bar{V}$ for this problem is the nxn matrix

$$\bar{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & & \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{bmatrix} \tag{196}$$

whose elements are given by

$$v_{ij} = q_{ij} \, E_{yi} \, E_{yj} \qquad (i,j = 1,n). \tag{197}$$

The formal solution to this problem is given by the following three equations:

$$\bar{C} = (\bar{A}^T \bullet \bar{V}^{-1} \bullet \bar{A})^{-1}, \tag{198}$$

$$\vec{p} = \bar{C} \bullet \bar{A}^T \bullet \bar{V}^{-1} \bullet \vec{y}, \tag{199}$$

$$\chi^2 = (\vec{y} - \bar{A} \bullet \vec{p})^T \bullet \bar{V}^{-1} \bullet (\vec{y} - \bar{A} \bullet \vec{p}). \tag{200}$$

χ^2 is distributed according to chi-square tables for (n-m) degrees of freedom. The covariance matrix $\bar{C}$ for the solution $\vec{p}$ is the mxm matrix

$$\bar{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & & & \\ c_{m1} & c_{m2} & \cdots & c_{mm} \end{bmatrix} . \tag{201}$$

The expansion coefficient errors $E_{p1},\ldots,E_{pm}$ $(\vec{E}_p)$ are given by

$$E_{pj} = (c_{jj})^{1/2} \quad (j = 1,m), \tag{202}$$

and the correlation matrix $\bar{D}$ for the coefficient errors is the mxm matrix

$$\bar{D} = \begin{bmatrix} 1 & d_{12} & \cdots & d_{1m} \\ d_{21} & 1 & \cdots & d_{2m} \\ \vdots & \vdots & & \vdots \\ d_{m1} & d_{m2} & \cdots & 1 \end{bmatrix} . \tag{203}$$

with elements given by

$$d_{ij} = c_{ij}/(c_{ii}\ c_{jj})^{1/2} \qquad (204)$$

$$(i,j = 1,m).$$

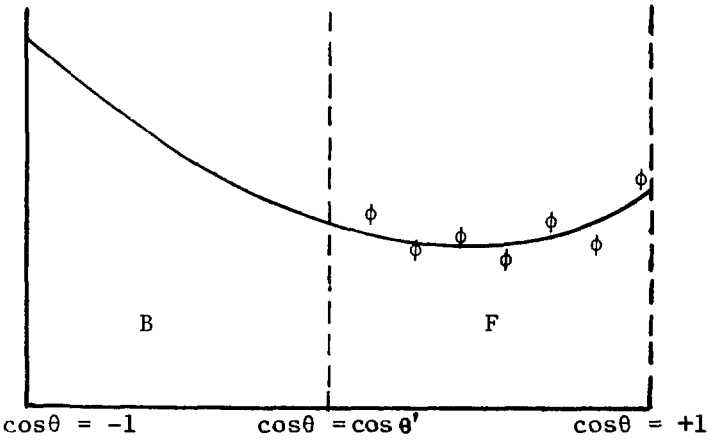
This elegant solution to a rather complex problem can be readily programmed for a digital computer-especially if explicit matrix manipulation routines are available. The procedure yields far more information about the solution than one derives from more commonly used least-squares formulations.

Example: Legendre Expansion Fit to Differential Plus Integral Data

The preceding problem was basically linear. We now consider a problem which is nonlinear. Again, we address the matter of fitting a Legendre polynomial expansion to angular distribution data. This time, the data points are explicitly given in the center of mass (c.m.) system, and we are faced with a situation whereby only forward-angle values are available. Back angle measurements were excluded by an experimental limitation. Experience has shown that forward angle data alone are often insufficient to enable a good fitted curve to be generated. Therefore, in this experiment, a determination was made in the laboratory system of the ratio of integrated forward events to integrated backward events. The problem is to devise a scheme for determining the "best" set of Legendre coefficients consistent with the available differential and integral information. The inclusion of integral information generates a nonlinear problem. For simplicity, we will consider a c.m. expansion with only three coefficients p_1, p_2 and p_3 $(\vec{p})$, then

$$y_c = Y(\theta) = p_1 + p_2\ P_1\ (\cos\ \theta) + p_3\ P_2(\cos\ \theta). \qquad (205)$$

Input data consists of n experimental angular distribution points $y_1,\ldots y_n$ $(\vec{y})$ in the c.m. and a measured laboratory forward-to-backward ratio R. The corresponding errors are $E_{y1},\ldots,E_{yn}$ $(\vec{E}_y)$ and E_R. The angular distribution points are assumed to be correlated as specified by an nxn matrix $\bar{Q}$ defined as per Eqs. (190) and (191). The forward-backward ratio R is assumed to be uncorrelated to the distribution points.



Eq. (205) provides an expression for calculating angular distribution points in the c.m. system in terms of the elements of $\vec{p}$. We need an expression which relates the calculated forward-backward ratio to the elements of $\vec{p}$. Since the calculated distribution applies to the c.m., and the forward-backward ratio is a laboratory quantity, some algebraic analysis is required to generate the required formula. This involves straightforward laboratory to c.m. conversions. We will simply give the result here. Define the following quantities: m_1 (incident particle mass), m_2 (target mass), m_3 (detected reaction product), m_4 (residual nucleus), Q (reaction Q-value), E_1 (incident particle energy in the laboratory), E_T (= E_1 + Q) and E_3 (energy of detected particle). Let

$$\beta = m_1 m_3 \, (E_1/E_T)/[(m_1 + m_2)(m_3 + m_4)], \tag{206}$$

$$\delta = m_2 m_4 \, (1 + m_1 Q/m_2 \, E_T)/[(m_1 + m_2)(m_3 + m_4)], \tag{207}$$

$$\theta' = \pi - \sin^{-1} \left[\left(\frac{\delta - \beta}{\delta}\right)^{1/2}\right]. \tag{208}$$

The angle θ' is the c.m. angle corresponding to 90° in the laboratory. The calculated forward yield F_c and backward yield B_c are determined from

$$F_c = \int_0^{\theta'} Y(\theta)\, d\theta = f_1 \, p_1 + f_2 \, p_2 + f_3 \, p_3, \tag{209}$$

$$B_c = \int_{\theta'}^{\pi} Y(\theta)\, d\theta = b_1 \, p_1 + b_2 \, p_2 + b_3 \, p_3. \tag{210}$$

The desired calculated ratio R_c is thus

$$R_c = F_c/B_c = \left(\frac{f_1\ p_1 + f_2\ p_2 + f_3\ p_3}{b_1\ p_1 + b_2\ p_2 + b_3\ p_3}\right) \tag{211}$$

where

$$f_1 = 1 - \cos\theta' \qquad b_1 = \cos\theta' + 1$$

$$f_2 = \frac{1}{2}\sin^2\theta' \qquad b_2 = -\frac{1}{2}\sin^2\theta' \tag{212}$$

$$f_3 = \frac{1}{2}\cos\theta' - \frac{1}{2}\cos^3\theta' \qquad b_3 = \frac{1}{2}\cos^3\theta' - \frac{1}{2}\cos\theta'$$

Clearly R_c is nonlinear in the expansion coefficients $\vec{p}$. We now proceed to linearize the problem and set up the equations for the least-squares analysis. We need an initial guess for the parameter set $\vec{p}$, called $\vec{p}_0$. This is not to hard to come by, especially if we remember that the scale (or normalization) can be estimated from the expression

$$\int_{4\pi} Y(\theta)\ d\Omega = 4\pi\ p_1. \tag{213}$$

Now define $\vec{x}$ according to

$$x_k = p_k - p_{0k} \quad (k = 1,2,3). \tag{214}$$

The vector $\vec{x}$ is the difference vector for the desired solution $\vec{p}$ and the initial guess (apriori) $\vec{p}_0$. Let $\vec{\xi}$ be the vector of n+1 experimental values:

$$\xi_i = y_i \ (i=1,n), \tag{215}$$

$$\xi_{n+1} = R. \tag{216}$$

Let $\vec{\xi}_0$ be the vector of calculated values based on apriori parameters $\vec{p}_0$:

$$\xi_{0i} = y_{c0i} = Y_0(\theta_i) \quad (i=1,n), \tag{217}$$

$$\xi_{0,n+1} = R_{c0}. \tag{218}$$

Let $\vec{z}$ be the difference vector defined by

$$z_i = \xi_i - \xi_{0i} \quad (i=1,\ n+1). \tag{219}$$

The first n linear equations required for the analysis are

$$y_i \approx y_{ci} = p_1 + p_2\ P_1\ (\cos\ \theta_i) + p_3\ P_2\ (\cos\ \theta_i), \quad (i=1,n), \tag{220}$$

and it can be shown readily that the set

$$z_i \approx x_1 + x_2\ P_1\ (\cos\ \theta_i) + x_3\ P_2(\cos\ \theta_i) \quad (i=1,n) \tag{221}$$

is completely equivalent. We now need an expression for z_{n+1}. R_c can be approximated by a truncated Taylor's series expansion in the coefficients $\vec{p}$ around the apriori $\vec{p}_0$. Based on Eq. (214), the formulas are:

$$R_c \approx R_{c0} + \left(\frac{\partial R_c}{\partial p_1}\right)_0 x_1 + \left(\frac{\partial R_c}{\partial p_2}\right)_0 x_2 + \left(\frac{\partial R_c}{\partial p_3}\right)_0 x_3, \tag{222}$$

$$\begin{cases} z_{n+1} \approx \left(\frac{\partial R_c}{\partial p_1}\right)_0 x_1 + \left(\frac{\partial R_c}{\partial p_2}\right)_0 x_2 + \left(\frac{\partial R_c}{\partial p_3}\right)_0 x_3, \\ \left(\frac{\partial R_c}{\partial p_k}\right)_0 = (B_{c0}\ f_k - F_{c0}\ b_k)/B_{c0}^2 \ (k = 1,2,3). \end{cases} \tag{223}$$

These results can be combined into the expression

$$\vec{z} \approx \bar{A} \bullet \vec{x} \tag{224}$$

where

$$\bar{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ & \vdots & \\ a_{n+1,1} & a_{n+1,2} & a_{n+1,3} \end{bmatrix}, \tag{225}$$

and

$$a_{i1} = 1,\ a_{i2} = P_1(\cos\ \theta_i),\ a_{i3} = P_2(\cos\ \theta_i), \quad (i=1,n) \tag{226}$$

$$a_{n+1,k} = \left(\frac{\partial R_c}{\partial p_k}\right)_0 \quad (k=1,2,3). \tag{227}$$

The use of the apriori parameter vector $\vec{p}_0$ has enabled us to generate a linear problem which relates differences $\vec{z}$ between the experimental quantities $\vec{\xi}$ and calculated apriori values $\vec{\xi}_0$, and differences $\vec{x}$ between the desired solution parameters $\vec{p}$ and the initial guess (apriori) set $\vec{p}_0$. No error is assumed for the apriori quantities, so the covariance matrix $\bar{V}$ for the difference quantities $\vec{z}$ can be derived readily from the experimental errors $\vec{E}_y$ and E_R for the input data, and the correlation matrix $\bar{Q}$ for the angular distribution point errors. So

$$\bar{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1,n+1} \\ v_{21} & v_{22} & \cdots & v_{2,n+1} \\ \vdots & \vdots & & \vdots \\ v_{n+1,1} & v_{n+1,2} & \cdots & v_{n+1,n+1} \end{bmatrix} \tag{228}$$

where

$$v_{ij} = q_{ij} E_{yi} E_{yj} \quad (i,j=1,n),$$

$$v_{n+1,n+1} = E_R^2, \tag{229}$$

$$v_{n+1,i} = v_{i,n+1} = 0 \quad (i=1,n).$$

The formal solution to the problem is provided by the formulas

$$\bar{C} = (\bar{A}^T \bullet \bar{V}^{-1} \bullet \bar{A})^{-1}, \tag{230}$$

$$\vec{x} = \bar{C} \bullet A^{-T} \bullet \bar{V}^{-1} \bullet \vec{z}, \tag{231}$$

$$\chi^2 = (\vec{z} - \bar{A} \bullet \vec{x})^T \bullet \bar{V}^{-1} \bullet (\vec{z} - \bar{A} \bullet \vec{x}), \tag{232}$$

$$\vec{p} = \vec{p}_0 + \vec{x}. \tag{233}$$

For many nonlinear problems, it is worthwhile to iterate at least once to see if χ^2 can be further minimized. This is accomplished readily by using the solution $\vec{p}$ from the first pass as the new apriori $\vec{p}_0$ for the second pass, and then repeating the analysis indicated above. Experience with problems of this nature has shown that if the initial $\vec{p}_0$ is anywhere near being reasonable, then convergence will be very rapid and little is gained by proceeding beyond two or three iterations. Examination of χ^2 (Eq. 232) at each stage of analysis provides the means for judging the progress toward convergence. The χ^2 series is derived from the formulas

$$\chi_o^2 = \vec{z}_o^T \bullet \bar{V}^{-1} \bullet \vec{z}_o \quad \text{(initial)}$$

$$\chi_i^2 = (\vec{z}_{i-1} - \bar{A} \bullet \vec{x}_i)^T \bullet \bar{V}^{-1} \bullet (\vec{z}_{i-1} - \bar{A} \bullet \vec{x}_i) \quad (i=1,2,\ldots), \tag{234}$$

where the subscript "i" indicates the particular iteration. A poor initial choice for the apriori parameters may cause the problem to converge toward a nonphysical solution, or oscillations may develop which prevent convergence of the series indicated in Eq. (234). When convergence is observed, the value of χ^2 provides a direct indication of the quality of the fit. If the final χ^2 indicates a very low probability in chi-square tables for the appropriate degrees of freedom (this example has n-2 degrees of freedom), then either the input data are inconsistent or the fitting procedure converged to a nonphysical solution. A plot of the results will usually indicate what has happened.

The matrix $\bar{C}$ from Eq. (230) has the form

$$\bar{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} , \tag{235}$$

and is the covariance matrix for the solution $\vec{x}$, and hence also for $\vec{p}$. The error vector $\vec{E}_p$ for the solution and its correlation matrix $\bar{D}$ can be derived from the formulas

$$E_{pk} = (c_{kk})^{1/2} \quad (k=1,2,3), \tag{236}$$

$$d_{ij} = c_{ij}/(E_{pi}\, E_{pj}) \quad (i,j=1,2,3), \tag{237}$$

where

$$\bar{D} = \begin{bmatrix} 1 & d_{12} & d_{13} \\ d_{21} & 1 & d_{23} \\ d_{31} & d_{32} & 1 \end{bmatrix} . \tag{238}$$

One final important point needs to be made before closing this section. In parameterizing data, such as was done in the preceding two examples, correlations will probably be introduced by the fitting process <u>even if the initial data are uncorrelated</u>. In parameterizing data, we convert from a "coordinate" system which may or may not be orthogonal to a totally

different system which probably will not be orthogonal. Therefore, one should use proper least-squares techniques which involve covariances, even in fitting uncorrelated data, if one wishes to gain a proper understanding of the uncertainties in the results.

XII. CONCLUSIONS

Covariance and least-squares methods are elegant and versatile analytic tools. The formalisms are simple and easy to use in principle. The usefulness of these methods for routine data fitting applications of the type discussed in Section XI would alone justify expending effort to learn the techniques.

The application of methods using covariances and least squares conditions poses a basic problem for the nuclear data community. The methods demand extensive input information which is often unavailable or can only be estimated crudely. The information which is provided as output may actually be more than is really required for applications. People who do reactor calculations routinely condense nuclear data by averaging and forming group sets. Is it reasonable to expect them to incorporate covariance information into their analyses, considering the already considerable complexity of their work even when detailed error information is not used? By analogy, a laboratory administrator would demonstrate poor judgment if he designated a large portion of his budget to the acquisition and maintenance of a costly piece of equipment which would most likely only be of limited value to the research program. A detailed error analysis of the type required to generate a realistic covariance matrix might take longer than the execution of an experiment and the subsequent data reduction. Some experiments are so complex that the experimenter would probably be forced to guess at the covariances involved. Might the time not be better spent doing another experiment? When one views the data base for applications from a broad perspective - e.g. glancing through data compilation plots such as those in BNL-325 (Ref. 14) - it is clear that there are still many deficiencies. Some of these may be important, and significant new information could be obtained by using new ideas and techniques in experiments which address these relatively untouched areas. Good judgment must be exercised. The problem reduces to the matter of assessing the need for detailed information versus the effort required to generate it. Evaluators are also faced with perplexing problems in applying these methods. If the experimenters do not provide the necessary information to generate covariance matrices, or if the experimenters do a poor job, then evaluators are forced to exercise subjective judgment in order to generate the needed information. The rigor of least-squares analyses should not mask the fact that subjective judgment is often required to prepare the input. The term "mechanical evaluation" is misleading since it only applies to part of the evaluation process.

One point is apparent. If covariance and least-squares methods are to be more widely used in nuclear data research, then considerable effort will have to be devoted to the development of approximation techniques. Experimenters and evaluators must learn how to identify the important factors in experiments. The key word is "sensitivity". For example, if it can be shown that the dominant correlations in a complicated experiment are introduced by only a few parameters, then the generation of an adequate covariance matrix may be feasible even for an experiment which generates a large number of data points.

The first proving ground for covariance and least-squares methods should be in the area of nuclear data standards. Nuclear data researchers are continuously reminded of the importance of accurate standards to their work.

If standard errors are small, then correlations between data sets based on these standards will be reduced and data evaluation and management will be less problematic. It seems quite reasonable to embark upon a program to evaluate the basic standards in a unified way as suggested by Poenitz (Ref. 8). This would enable us to make the best possible use of the existing data base, and would indicate more precisely what new standard measurements are needed to achieve realistic accuracy goals, and what accuracies will be required for these measurements if they are to make an impact.

Beyond the field of standards, the need for covariance information must be examined on a case-by-case basis. With a few possible exceptions, e.g., the dosimetry field, it appears that the need for detailed covariance information is not well established.

ACKNOWLEDGEMENTS

This work was undertaken as the result of a research problem suggested by H. Liskien, Central Bureau for Nuclear Measurements, Geel, Belgium. His thoughtful review of much of the material contained in this report is appreciated. This work also benefitted from comments made by W. P. Poenitz, Argonne National Laboratory, Illinois, U.S.A.

REFERENCES

1. Yardley Beers, Introduction to the Theory of Error, Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, U.S.A., (1966).

2. I. S. Sokolnikoff and R. M. Redheffer, Mathematics of Physics and Modern Engineering, 2nd Edition, McGraw-Hill Publishing Company, New York, N.Y., U.S.A., (1966).

3. R. W. Peele, "Requirements on Experiment Reporting to Meet Evaluation Needs," Proceedings of the NEANDC/NEACRP Specialists Meeting on Fast Neutron Fission Cross Sections of U-233, U-235, U-238, and Pu-239," June 28-30, 1976 at Argonne National Laboratory, edited by W. P. Poenitz and A. B. Smith, p. 421, Report Nos. ANL-76-90 or NEANDC (US)-199/L (1976).

4. F. G. Perey, "Contributions to Few-Channel Spectrum Unfolding," Report No. ORNL/TM-6267, Oak Ridge National Laboratory (1978).

5. F. G. Perey, "Covariance Matrices of Experimental Data," Proceedings of an International Conference on Neutron Physics and Nuclear Data for Reactors, Harwell, United Kingdom, p. 104 (September 1978).

6. W. Mannhart and F. G. Perey, "Covariance Matrices of Cf-252 Spectrum Averaged Cross Sections", Proc. Third ASTM-EURATOM Symposium on Reactor Dosimetry, Ispra, Italy, Oct. 1-5, 1979, Report EUR-6813, Vol. II, p. 1016 (1980).

7. W. Mannhart, "A Small Guide to Generating Covariances of Experimental Data," Report No. PTB-FMRB-84, Physikalisch-Technische Bundesanstalt, Braunschweig (June 1981).

8. W. P. Poenitz, "Data Interpretation, Objective Evaluation Procedures and Mathematical Techniques for the Evaluation of Energy-Dependent Ratio, Shape and Cross Section Data," Proceedings of the Conference on Nuclear Data Evaluation Methods and Procedures, September 22-25, 1980 at Brookhaven National Laboratory, edited by B. A. Magurno and S. Pearlstein, Report Nos. BNL-NCS-51363 or NEANDC (US)-209, Vol. I, p. 249 (March 1981).

9. Philip R. Bevington, Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill Book Company, New York, N.Y., U.S.A. (1969).

10. Standards File ENDF/B-V, National Nuclear Data Center, Brookhaven National Laboratory, Upton, N.Y. 11973, U.S.A.

11. ENDF-102, Data Formats and Procedures for the Evaluated Nuclear Data File, ENDF, Revised by R. Kinsey, Report BNL-NCS-50496, Brookhaven National Laboratory, Upton, New York 11973, U.S.A. (Oct. 1979).

12. J. J. Schmidt, "Summary Review of Second-Day Sessions", ibid. Ref. 8, p. 501.

13. S. Tagesen, H. Vonach and B. Strohmaier, "Evaluations of the Cross Sections for the Reaction $^{24}Mg(n,p)^{24}Na$, $^{64}Zn(n,p)^{64}Cu$, $^{63}Cu(n,2n)^{62}Cu$ and $^{90}Zr(n,2n)^{89}Zr$, Physics Data 13-1 (1979).

14. D. I. Garber and R. R. Kinsey, "Neutron Cross Sections, Vol. II, Curves," Report BNL-325 (3rd Edition) - Vol. II, National Nuclear Data Center, Brookhaven National Laboaratory, N.Y., U.S.A. (1976).