

NAS-NS
3109

**National
Academy
of
Sciences**

National Research Council

NUCLEAR SCIENCE SERIES
Radiochemical Techniques

**Processing of
Counting Data**

U.S.
Atomic
Energy
Commission

COMMITTEE ON NUCLEAR SCIENCE

Chairman
R. G. Bunch

Chairman
W. J. Bunker

Secretary
David Clark
National Research Council

E. C. Anderson
Los Alamos Scientific Laboratory

Terry B. Marion
University of Maryland

N. E. Ballou
U. S. Naval Radiological Defense
Laboratory

R. L. Platman
Argonne National Laboratory

Martin J. Berger
National Bureau of Standards

Ernest C. Pollard
Pennsylvania State University

C. J. Borkowski
Oak Ridge National Laboratory

Katharine Way
Oak Ridge National Laboratory

Herbert Goldstein
Columbia University

George W. Wetherill
University of California

Bernd Kahn
Fairchild Engineering Center

Marvin E. Wyman
University of Illinois

LIAISON MEMBERS

Harold Glaser
Office of Naval Research

William S. Rodney
National Science Foundation

George A. Kolstad
Atomic Energy Commission

SUBCOMMITTEE ON RADIOCHEMISTRY

Chairman
Nathan E. Ballou
U. S. Naval Radiological Defense
Laboratory

Julian M. Nielsen
Battelle Pacific Northwest

G. R. Choppin
Florida State University

G. D. Eckart
Oak Ridge National Laboratory

Herbert M. Clark
Rensselaer Polytechnic Institute

R. P. Steinhilber
Argonne National Laboratory

Richard M. Diamond
Lawrence Radiation Laboratory

D. N. Sunderman
Battelle Memorial Institute

Jerome Hudis
Brookhaven National Laboratory

John W. Winchester
Massachusetts Institute of Technology

Jerome D. Knigh
Los Alamos Scientific Laboratory

R. P. Schuman
Atomic Energy Division
Phillips Petroleum Company (Ciano Falls)

W. J. Bunker
Lawrence Radiation Laboratory

Processing of Counting Data

P. C. Stevenson

Lawrence Radiation Laboratory, University of California
Livermore, California

September 24, 1965

I. INTRODUCTION

1. General Contents

Counting data is the term which I propose to use loosely to describe the great variety of results of measurements on amount and kind of radiation emanating from a radioactive sample. In general, counting data are taken as a function of time after a certain preassigned zero time in terms of events per unit time. The instruments used to accumulate such data are many and various. They have been described in a number of publications (for a simple summary see another monograph of this series¹). The interpretation and use of such data involve a number of complications, many of a statistical nature. This monograph is intended to be a handy reference manual of various ways of treating and interpreting data. It is also intended to serve as a reminder of some neglected aspects of the treatment of data and as a collection of information for those who are not familiar with the field.

Chapter II, Mathematics of Radioactive Processes, is concerned with the phenomenology of radiation: with its decay with time; with the formation of radioactive species by nuclear reactions; with consecutive processes (consecutive formations or consecutive decays) and competitive processes; and with the relationship of the parameters of nuclear reactions and nuclear processes to the raw data.

Chapter III on Mathematical Techniques gives a very elementary outline of the minimum necessary mathematics, matrix algebra, elementary statistical mathematics, a description of the least-squares method of curve fitting, and of the χ^2 test as a criterion of goodness of fit. There is also a

¹G. D. O'Kelley, "Detection and Measurement of Nuclear Radiation,"
NAS-NS-3105.

FOREWORD

The Subcommittee on Radiochemistry is one of a number of subcommittees working under the Committee on Nuclear Science within the National Academy of Sciences - National Research Council. Its members represent government, industrial, and university laboratories in the areas of radiochemistry and nuclear chemistry. Support for the activities of this and other subcommittees of the Committee on Nuclear Science is provided by a grant from the National Science Foundation.

The Subcommittee has concerned itself with preparation of publications, encouraging and supporting activities in nuclear education, sponsoring symposia on selected current topics in radiochemistry and nuclear chemistry, and investigating special problems as they arise. A series of monographs on the radiochemistry of essentially all the elements and on radiochemical techniques is being published. Initiation and encouragement of publication of articles on nuclear education in various subject areas of chemistry have occurred, and development and improvement of certain educational activities (e.g., laboratory and demonstration experiments with radioactivity) have been encouraged and assisted. Radioactive contamination of reagents and materials has been investigated and specific recommendations made.

This series of monographs has resulted from the need for comprehensive compilations of radiochemical and nuclear chemical information. Each monograph collects in one volume the pertinent information required for radiochemical work with an individual element or with a specialized technique. The U. S. Atomic Energy Commission has sponsored the printing of the series.

Comments and suggestions for further publications and activities of value to persons working with radioactivity are welcomed by the Subcommittee.

N. E. Ballou, Chairman
Subcommittee on Radiochemistry

PREFACE

In this monograph I have attempted to collect a number of problems, and their solutions, which are of importance to radiochemists who are involved with the collection, reduction, and interpretation of various kinds of data arising from measurements using counting equipment.

In cases where the treatment of a problem has been taken from another author, I have endeavored to give full credit to the original source; errors in transcription or rephrasing are entirely my own. When no credit is given, I accept responsibility for the treatment and for the unhappiness which some of my procedures will cause more qualified statisticians. My only defense is that I am attempting to write in a language which will be comprehensible to chemists.

This monograph has been prepared at the request of the Subcommittee on Radiochemistry of the Committee on Nuclear Science, Division of Physical Sciences, National Academy of Sciences - National Research Council, as one of the series of monographs on the techniques of radiochemistry which the Subcommittee is sponsoring. I would like to express my appreciation to the members of the Subcommittee for their encouragement; to Drs. E. Kaplan and H. B. Levy for their helpful discussions of the problems which arose during preparation of the manuscript; to the secretarial staff of the Radiochemistry Division and to the Technical Information Division of the Lawrence Radiation Laboratory for their assistance in preparing the manuscript.

This work was performed under the auspices of the U. S. Atomic Energy Commission.

P. C. Stevenson

Lawrence Radiation Laboratory
University of California
Livermore, California

TABLE OF CONTENTS

	<u>Page No.</u>
Foreword	-v-
I. Introduction	1
1. General contents	1
II. Mathematics of Radioactive Processes	3
1. Assumptions	3
2. Fundamental law of radioactive decay	3
3. Formation of species in nuclear reactions	5
4. Consecutive disintegrations	10
5. Use of daughter activities to measure nuclear reactions	13
6. Branching decay	19
7. Decay chains with more than two consecutive decays	21
III. Mathematical Techniques	25
1. Nature of the problem	25
2. Matrix algebra	26
3. Statistical mathematics	31
4. Averaging and sampling statistics	39
5. The least-squares solution	47
IV. Statistics of Counting Data	64
1. Nature of the problems	64
2. Distribution functions of counting results	64
3. Distribution of expected values for a given observed count - small sampling	68
4. Distribution of expected values for a given observed count - large sampling	69
5. Correction for counter background	76
6. Errors of extrapolation	79
7. Statistical reliability of measurements on nuclear reactions	82
V. The Resolution of Decay Data	83
1. Nature of the problem	83
2. The least-squares solution; weighting of the data	84
3. Estimation of variance due to instrumental variation	87

TABLE OF CONTENTS (Continued)

	<u>Page No.</u>
4. Effect of input parameters on resolution	89
5. Criteria for the rejection of data	92
6. Avoidance of computer overflow	94
7. Analysis of replicate samples; analysis of variance	95
VI. The Identification of Components and Measurement of Half-Lives or Decay Constants	101
1. Nature of the problems	101
2. Single-component system of short or moderate half- life	102
3. Single-component system with long half-life	103
4. Multicomponent systems - detection	105
5. Known multicomponent systems - evaluation of λ	107
6. Unknown multicomponent systems	109
VII. Dead Time and Coincidence Considerations	110
1. Nature of the problem	110
2. Effect of dead time on observed counting rate	112
3. Measurement of dead time	114
4. Scaling circuits - avoiding losses in slow components	117
5. Measurement of coincident events	120
6. Variance of coincident event rate	123
7. The effect of coincident events on discriminated counts	124
8. Measurement of absolute counting efficiency by coincidence techniques	127
VIII. Glossary of Symbols	129
IX. Index	131

Processing of Counting Data

P. C. Stevenson

UNIVERSITY OF CALIFORNIA
Lawrence Radiation Laboratory
Livermore, California

Issued: May 1966

Subcommittee on Radiochemistry
National Academy of Sciences—National Research Council

Printed in USA. Price \$1.75. Available from the Clearinghouse for Federal Scientific and Technical Information, National Bureau of Standards, U. S. Department of Commerce, Springfield, Virginia.

description of a so-called "F" test which is used to aid in making a decision whether the form of equation being fitted to the data is proper.

Chapter IV describes the statistical problems particularly associated with counting data. Variance of a single counting measurement is discussed in some detail. The effect of making a measurement involving small sampling, which is the usual method of procedure, is described and is followed by a description of the alterations in the treatment which should be used when large sampling or low count rates are involved. Statistical problems raised by correction for background are discussed, as are errors of extrapolation.

Chapter V is a detailed description of resolution methods applied to decay data, which involve separating a given set of data into components each attributable to a particular radioactive species. Proper weighting for least-squares analysis is discussed. There is a description of sources of variance and of the effect of input parameters such as half-lives. Criteria for data rejection are discussed, as are analyses of replicate samples.

Chapter VI is a description of the problems involved in identification of the components of the system and measurement of their decay characteristics; in particular, their half-lives. The problem of purity is discussed. The problem of measurement of half-life of single-component systems of short half-life and of long half-life is considered in some detail. Another problem considered in this section is the adjustment of half-lives with the best possible precision in a known multicomponent system where the half-lives are not well known. This section also discusses the detection of multicomponent systems, the determination of the number of components in a system, and the measurement of the half-lives where the system components are completely unknown.

The last section discusses problems raised by the existence of dead time and coincidence phenomena. Dead time losses in counters are discussed. Live timers, which are mechanical methods of compensating for dead time loss, are mentioned. Regularizing action of scaling circuits is discussed in some detail. Measurement of coincident events and the problems raised by chance coincidences are discussed and the effect of coincidences on discriminated counts are considered. Finally the measurement of absolute counting efficiency by coincidence techniques is mentioned briefly.

II. MATHEMATICS OF RADIOACTIVE PROCESSES

1. Assumptions

It will be assumed in this section that: (a) the intensity of radiation emerging from a sample of radioactive material can be measured with high precision and with known, or at least reproducible, efficiency; (b) that all atoms of a single radioactive species are identical in their nuclear properties; (c) that the intensity of the radiation originating in the sample is due to the nuclear properties of the species present; and (d) that each radioactive event signals the disintegration of one atom of the radioactive species and the formation of one atom of a new species. Assumptions (b), (c), and (d) have been verified experimentally; assumption (a) is adopted in this section for simplicity to represent an ideal case and will be discarded in later sections.

2. Fundamental Law of Radioactive Decay

The assumptions listed above lead to the obvious conclusion that the amount of radiation originating in a sample is directly proportional to the amount of radioactive material present (although the constant of proportionality varies from species to species). For a single species, if x is the number of atoms present at time t , the intensity of radiation in events per unit time is given by the rate at which the active species is disappearing (assumption (d)), so that the intensity of radiation is

$$-\frac{dx}{dt} = +\lambda x \quad (\text{II.1})$$

where λ is the constant of proportionality characteristic of the species. λ is referred to as the decay constant of the species. The negative sign in the equation arises from assumption (d) above, that one radioactive event corresponds to the disappearance of one atom of the species. (It should be noted that we are ignoring the discrete nature of atomic matter in this section and treating x as a continuous variable. This is consistent with our first assumption.) Equation II.1 represents one statement of the fundamental law of radioactive decay. It can be integrated immediately to give

$$x = x_0 e^{-\lambda(t-t_0)} \quad (\text{II.2})$$

where x has the value x_0 at the arbitrarily chosen time t_0 . The product λx can be seen from equation II.1 to represent the disintegration rate of the species; if we measure the radiation from the sample with an instrument having an efficiency η for the characteristic radiations of our species, the observed "counting rate" will be given by $\eta\lambda x$. Both x and $\eta\lambda x$ can be seen from equation II.2 to decrease exponentially with time.

A convenient measurement for purposes of discussion and observation is the time interval required for the activity of the species in the sample to decrease to one-half of its initial value. This interval $T_{1/2}$ is referred to as the half-life, half-period, or simply period. Its value in terms of λ can be derived by setting x equal to $x_0/2$ in equation II.2, so that

$$e^{-\lambda T_{1/2}} = \frac{1}{2} \quad (\text{II.3})$$

and therefore

$$T_{1/2} = \frac{\ln 2}{\lambda} . \quad (\text{II.4})$$

$T_{1/2}$ is expressed in time units and may be specified in units of any convenient size. λ is expressed in units of reciprocal time; the nature of the units used is of particular significance when calculating disintegration rates from numbers of atoms or the reverse, since λ and λx must be expressed in the same units.

If more than one species are present, the effects are additive, that is

$$\text{Counting rate} = \sum_i \eta_i \lambda_i x_i = \sum_i \eta_i \lambda_i x_{i0} e^{-\lambda_i t(t-t_0)} \quad (\text{II.5})$$

In such a system, each separate species is referred to as a component. The process of identifying the various components in an experimentally obtained set of data and of separating the effect due to each is called resolving the decay data. Information to be obtained in an experiment on radioactive material is usually gained from a measurement of the amount, the half-life, or the nature of the radiations of one or more components separately; it is with the resolution of multicomponent systems that the greater part of this paper will deal.

The most elementary technique for resolving data on radioactive decay depends on the fact that in any multicomponent mixture, there is usually one component much longer lived than any of the others. At late times, therefore, the activity still remaining in the system is likely to be due to a single long-lived component. If the activity of the system at various times is plotted as ordinate against time as abscissa on semi-logarithmic graph paper (see Fig. II.1), and a smooth curve drawn through the points, the last portion of the curve obtained will frequently be a good approximation to a straight line. (If only one component is present, the entire curve will be a straight line.) This straight line can be extrapolated back to the times of the early observations and the contribution of the longest lived species subtracted from each gross amount. The differences are then replotted and the procedure repeated for the next-longest lived species. This technique is simple, but not very precise and inevitably prone to subjective judgments.

The straight lines obtained from the above-described plots may be used to get an estimate of the half-lives of the components by observing the time required for the amount of each component to decrease by exactly a factor of 2, since the linear nature of the plots makes interpolation to the desired activity level particularly simple. Somewhat better precision may be obtained by observing the time required for the activity to decrease by an integral power of 2, and dividing that time by the power used.

3. Formation of Species in Nuclear Reactions²

Radioactive species are formed either by nuclear reactions between nuclei and subatomic particles or photons or by spontaneous radioactive decay of other radioactive parent species. In this section we shall consider the number of atoms x of a radioactive species present in a sample of material exposed to conditions leading to the formation of the radioactive species, e. g., in a cyclotron beam or a flux of neutrons in a reactor.

Let the flux of particles at each atom in the target material be ϕ . The rate of the nuclear reaction will be proportional to the flux integrated over the entire sample. If the flux is uniform, this is simply the product of the flux ϕ by the number n of atoms in the sample; if the flux is limited to a beam passing through a sample of uniform thickness, it is the product of the total beam current by the thickness of the sample in atoms per unit area. In actual practice the flux is usually not uniform; gradients in the flux in a sample of finite size may arise either from the gradient in space of the flux into which the sample is inserted or from the attenuation of the flux by the sample itself. One must then integrate the flux at each infinitesimal volume element, times the number of atoms of target in that volume element, over the volume of the entire target. In

² See Rubinson, J. Chem. Phys. 17, 542 (1949).

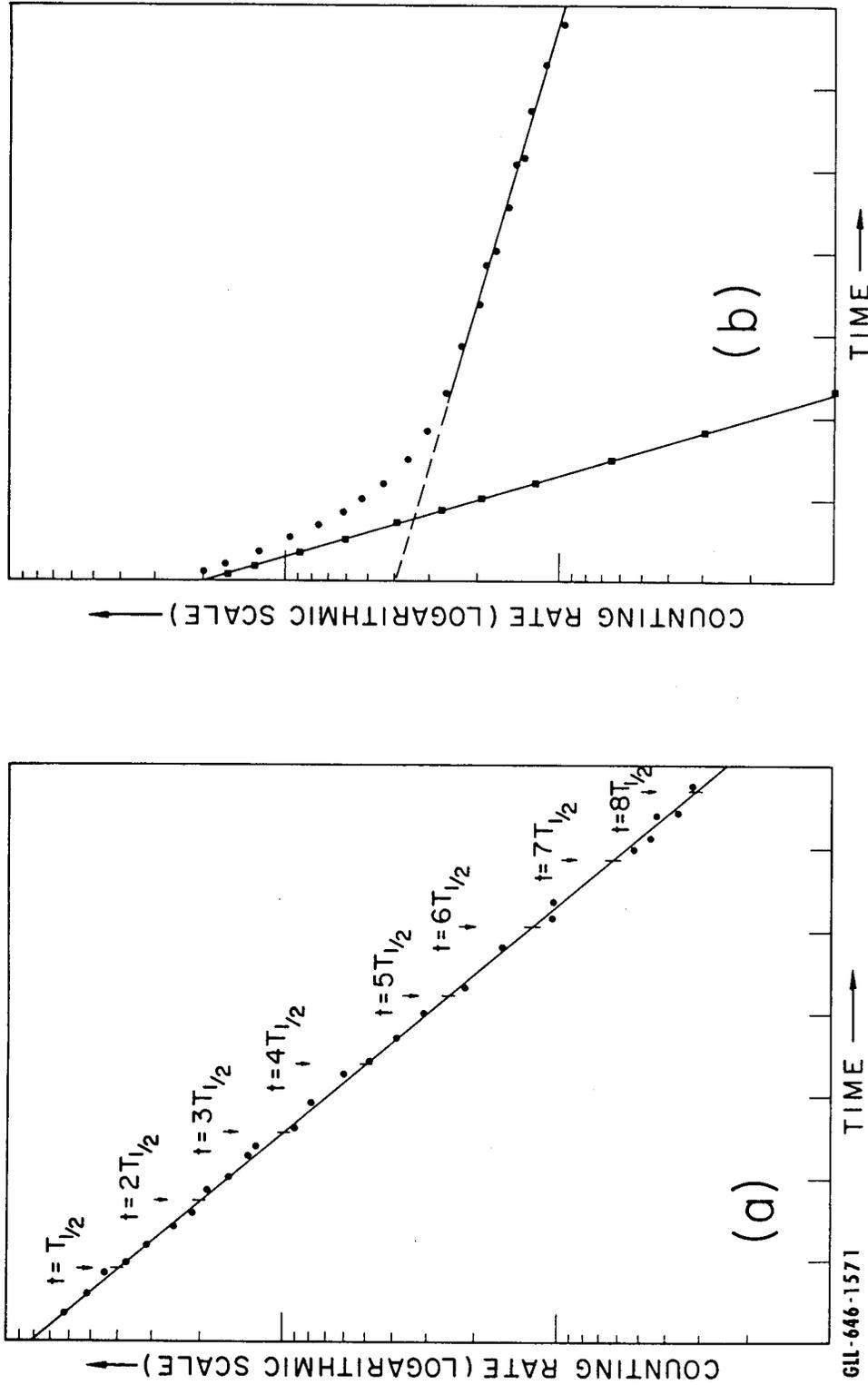


Fig. 11.1. (a) Activity of a single-component radioactive sample plotted against time. Arrows mark the points on the line at which the activity has decreased by successive factors of 2. Time interval between arrows is one half-life. The best graphical estimate of the half-life is obtained by dividing the time interval between the zero time and the time of the n th arrow by n , where n is as great as possible. No real advantage is gained by extrapolating the curve beyond the measured data, although it is convenient to extrapolate to the next even half-life.

(b) Activity of a two-component radioactive sample plotted against time. Circles are measured data points. Squares are obtained by subtracting the value of the extrapolated long-lived activity (dashed curve) from each measurement. Half-lives of the two species differ by approximately a factor of 10.

GIL-646-1571

either case, we shall use the symbolic product $n\phi$ to represent this integral. We shall define a reaction cross section σ to be the constant of proportionality which converts the product $n\phi$ to the actual rate of the nuclear reaction. It is apparent that if we consider the reaction to be occurring between an assembly of flying particles and a collection of atoms, then σ is simply the "target area" presented to the particles by a single atom.

The net rate of formation, however, contains another term, since the species being formed is radioactive and hence subject to the laws governing radioactive decay. The expression for the net rate of formation of the species is therefore given by

$$\frac{dx}{dt} = n\phi\sigma - \lambda x \quad (\text{II.6a})$$

where the second term on the right arises from the decay laws, expressed in the form of equation II.1. In most cases, $n\phi$ may be considered constant, since most methods of inducing nuclear reactions are extremely inefficient in terms of conversion of the target material. If this assumption is made, it is obvious that, whatever the starting conditions, the system will tend in the long run to an equilibrium value of x , since for small x the derivative is positive so that x is increasing, while for large x the derivative is negative so that x is decreasing. When the system has its equilibrium value of x , it is said to be saturated.

If $n\phi$ does not change with time, equation II.6a can readily be integrated by rearranging the terms and applying an integrating factor, as follows:

$$\frac{dx}{dt} + \lambda x = n\phi\sigma, \quad (\text{II.6b})$$

$$e^{\lambda t} \left(\frac{dx}{dt} + \lambda x \right) = n\phi\sigma e^{\lambda t}, \quad (\text{II.6c})$$

$$\frac{d}{dt} (x e^{\lambda t}) = n\phi\sigma e^{\lambda t}, \quad (\text{II.6d})$$

$$x e^{\lambda t} - x_0 e^{\lambda t_0} = \frac{n\phi\sigma}{\lambda} (e^{\lambda t} - e^{\lambda t_0}), \quad (\text{II.6e})$$

and finally,

$$x = x_0 e^{-\lambda(t-t_0)} + \frac{n\phi\sigma}{\lambda} \left[1 - e^{-\lambda(t-t_0)} \right]. \quad (\text{II.7})$$

In this equation the first term represents the behavior of the radioactive material present at the time t_0 , decaying away according to equation II.2, while the second term represents the buildup toward equilibrium of the material formed by nuclear reactions in the time following t_0 .

Equation II.7 only applies as long as the reaction is continuing; when the cyclotron beam is turned off or the sample is removed from the reactor, the active material present begins to decay away, following equation II.2 with t_0 for that equation being taken as the time of the end of the irradiation. The counting rate measured at some later time can always be extrapolated back to the end of the irradiation if the half-life is known; in the subsequent discussion it will be assumed that this has been done.

If $n\phi$ is not constant with time, equation II.7 cannot be obtained in analytic form unless the dependence of $n\phi$ on t is expressed analytically. Treatment of equation II.6a proceeds as before up to the integration step II.6e, which then becomes

$$d(xe^{\lambda t}) = \sigma e^{\lambda t} n\phi(t) dt \quad (\text{II.8})$$

leading to

$$x = x_0 e^{-\lambda(t-t_0)} + \sigma e^{-\lambda t} \int_{t_0}^t n\phi(t) e^{\lambda t} dt \quad (\text{II.9})$$

which usually must be evaluated numerically. In what follows we shall assume that flux at any point in the target is constant with time.

In a typical experiment, inactive material is irradiated at a constant rate for a defined length of time in a constant flux; x_0 is zero and equation II.7 can be used to give a measurement of ϕ if σ is known, or of σ if ϕ is known:

$$x = \frac{\phi n \sigma}{\lambda} \left[1 - e^{-\lambda(t-t_0)} \right], \quad (\text{II.10a})$$

$$\sigma = \frac{1}{\phi n} \frac{\lambda x}{\left[1 - e^{-\lambda(t-t_0)} \right]} = \frac{1}{\eta \phi n} \frac{\text{counting rate}}{\left[1 - e^{-\lambda(t-t_0)} \right]}, \quad (\text{II.10b})$$

$$\phi = \frac{1}{\sigma n} \frac{\lambda x}{\left[1 - e^{-\lambda(t-t_0)} \right]} = \frac{1}{\eta \sigma n} \frac{\text{counting rate}}{\left[1 - e^{-\lambda(t-t_0)} \right]}. \quad (\text{II.10c})$$

In either case the detector efficiency η must be known, which is frequently not the case; if η is not known relative numbers can be obtained, since η does not vary from experiment to experiment if the equipment is reliable.

If a calibration experiment is made with a known amount of material in a uniform flux, a useful "figure of merit" $\eta\phi\sigma$ may be obtained:

$$\eta\phi\sigma = \frac{\text{counting rate}}{n \left[1 - e^{-\lambda(t-t_0)} \right]} = a. \quad (\text{II.10d})$$

This figure of merit may then be applied to irradiations of unknown materials in the same flux to determine the amount of the target species in the unknown sample, since

$$n = \frac{\text{counting rate}}{a \left[1 - e^{-\lambda(t-t_0)} \right]}. \quad (\text{II.10e})$$

This is the basis of the very sensitive analytical method known as activation analysis. Care must be taken to reproduce the flux exactly; in practice, the calibration sample and the unknown are usually run simultaneously.

Returning to equation II.6a, it is sometimes necessary to modify our simplifying assumption. Development of high-flux reactors and other devices which are relatively efficient, though rather spectacularly short-lived, render it necessary to consider the case in which the amount of target material is appreciably depleted during the course of the experiment. For simplicity, we shall assume that the flux in this case is constant and uniform. We then have, since formation of one atom of x implies destruction of one atom of n,

$$\frac{dn}{dt} = -\phi n \sigma \quad (\text{II.11})$$

so that

$$n = n_0 e^{-\phi\sigma(t-t_0)} \quad (\text{II.12})$$

which may be substituted into equation II.6a. Following a method of integration analogous to that used on equation II.6a, we finally get

$$x = x_0 e^{-\lambda(t-t_0)} + \frac{\phi n_0 \sigma e^{-\phi\sigma(t-t_0)}}{\lambda - \phi\sigma} \left[1 - e^{-(\lambda - \phi\sigma)(t-t_0)} \right]. \quad (\text{II.13})$$

After very long irradiation times, this expression tends toward zero, owing to complete conversion of the target material followed by radioactive decay of the product. If we maximize x with respect to time, we can easily see that the greatest amount of product is present in the target at a time given by (assuming $x_0 = 0$)

$$t_m - t_0 = \frac{\ln(\lambda/\phi\sigma)}{\lambda - \phi\sigma}. \quad (\text{II.14})$$

Finally, if the product itself reacts with the flux, as it usually does, we must modify equation II.6a still further by addition of a term $-\phi\sigma_x x$ where σ_x is the reaction cross section of the species x with the flux. It is obvious that the only difference made in the mathematical form of the equation is the replacement of the constant λ with a new constant $\lambda + \phi\sigma_x$. The integrated form of the equation is then readily obtained by substitution in equation II.13, and the time of maximum production by substitution in equation II.14, giving

$$x = x_0 e^{-(\lambda + \phi\sigma_x)(t-t_0)} + \frac{\phi n_0 \sigma e^{-\phi\sigma(t-t_0)}}{\lambda + \phi\sigma_x - \phi\sigma} \left[1 - e^{-(\lambda + \phi\sigma_x - \phi\sigma)(t-t_0)} \right], \quad (\text{II.15})$$

$$t_m - t_0 = \frac{\ln\left(\frac{\lambda + \phi\sigma_x}{\phi\sigma}\right)}{\lambda + \phi\sigma_x - \phi\sigma}. \quad (\text{II.16})$$

4. Consecutive Disintegrations

It frequently happens that the disintegration product of a radioactive species is itself radioactive. The mathematical analysis of the system is quite straightforward; it is only necessary to keep in mind, first that the fundamental law of decay applies to all active species, and second that disintegration of one atom of the "parent" species corresponds to formation of one atom of the "daughter" species.

Let us consider a simple system, a species 1 decaying to a species 2 which decays to a stable end product. We have for the parent species

$$x_1 = x_{10} e^{-\lambda_1(t-t_0)}, \quad (\text{II.17a})$$

$$\eta_1 \lambda_1 x_1 = \eta_1 \lambda_1 x_{10} e^{-\lambda_1(t-t_0)} \quad (\text{II.17b})$$

for the amount and counting rate, respectively, of the parent species. The net rate of formation of the daughter species is given by

$$\frac{dx_2}{dt} = -\lambda_2 x_2 + \lambda_1 x_1 = -\lambda_2 x_2 + \lambda_1 x_{10} e^{-\lambda_1(t-t_0)}. \quad (\text{II.18})$$

Equation II.18 may be integrated in a manner analogous to that used for equation II.6a to give

$$x_2 = x_{20} e^{-\lambda_2(t-t_0)} + \frac{\lambda_1}{\lambda_2 - \lambda_1} x_{10} \left[e^{-\lambda_1(t-t_0)} - e^{-\lambda_2(t-t_0)} \right], \quad (\text{II.19a})$$

$$\eta_2 \lambda_2 x_2 = \eta_2 \lambda_2 x_{20} e^{-\lambda_2(t-t_0)} + \frac{\eta_2 \lambda_1 \lambda_2}{\lambda_2 - \lambda_1} x_{10} \left[e^{-\lambda_1(t-t_0)} - e^{-\lambda_2(t-t_0)} \right]. \quad (\text{II.19b})$$

(In the highly unlikely event that $\lambda_1 = \lambda_2 = \lambda$, it is a simple matter to show that

$$x_2 = x_{20} e^{-\lambda(t-t_0)} + \lambda x_{10} (t - t_0) e^{-\lambda(t-t_0)} \quad (\text{II.19c})$$

either by rederiving equation II.19a or by extracting the factor $e^{-\lambda_1(t-t_0)}$ from the parenthesis in equation II.19a to leave a term $\left[1 - e^{-(\lambda_2 - \lambda_1)(t-t_0)} \right]$. This term approaches $(\lambda_2 - \lambda_1)(t - t_0)$ as a limit as $(\lambda_2 - \lambda_1)$ approaches zero. This case will not be considered further.)

The total counting rate of the sample is obtained by adding the rates given by equations II.17b and II.19b to give

$$\begin{aligned} \text{Counting rate} = & \eta_1 \lambda_1 x_{10} e^{-\lambda_1(t-t_0)} + \eta_2 \lambda_2 x_{20} e^{-\lambda_2(t-t_0)} \\ & + \frac{\eta_2 \lambda_1 \lambda_2}{\lambda_2 - \lambda_1} x_{10} \left[e^{-\lambda_1(t-t_0)} - e^{-\lambda_2(t-t_0)} \right]. \quad (\text{II.20}) \end{aligned}$$

Combining terms with like exponential factors, we get

$$\begin{aligned} \text{Counting rate} = & \lambda_1 x_{10} e^{-\lambda_1(t-t_0)} \left[\frac{\lambda_2 (\eta_1 + \eta_2) - \lambda_1 \eta_1}{\lambda_2 - \lambda_1} \right] \\ & + \eta_2 \lambda_2 e^{-\lambda_2(t-t_0)} \left[\frac{x_{20} (\lambda_2 - \lambda_1) - \lambda_1 x_{10}}{\lambda_2 - \lambda_1} \right]. \end{aligned} \quad (\text{II.21})$$

It can easily be seen from equation II.21 that the decay of a parent-daughter system resembles that of any two-component system, and can be resolved on that basis. The half-lives of the apparent components are the actual half-lives of the parent and the daughter. However, when performing an experiment in which the product of interest gives rise to an active daughter, the presence of the daughter activity may lead to serious error unless an appropriate correction factor is applied. Let us assume that an irradiation is made and that a product (species 1) is separated which has an active daughter (species 2). The product is purified, the decay of the sample is followed for a time and the data resolved into its two components. We then attempt to apply equations II.10b or II.10c to the system, using the resolved counting rate of the component having the half-life of species 1. At the time of purification of the parent species, equation II.21 gives

$$\text{Counting rate (1)} = \lambda_1 x_1 \left[\frac{\lambda_2 (\eta_1 + \eta_2) - \lambda_1 \eta_1}{\lambda_2 - \lambda_1} \right] \quad (\text{II.22})$$

while in the absence of the daughter the counting rate at the same time would be

$$\text{Counting rate (no daughter)} = \lambda_1 x_1 \eta_1. \quad (\text{II.23})$$

The ratio of the observed to the desired rate is the growth factor g , which is independent of time (since both numerator and denominator in the ratio have the same time dependence) and consequently applies at the end of the irradiation. It is given by

$$g = 1 + \left(\frac{\lambda_2}{\lambda_2 - \lambda_1} \right) \left(\frac{\eta_2}{\eta_1} \right). \quad (\text{II.24})$$

Equations II.10b and II.10c apply to the counting rate in the absence of daughter activity; therefore the observed resolved counting rate must be divided by g before being inserted in these equations.

It will be observed that, unless η_2 is very small, as λ_2 approaches λ_1 the value of g becomes very large either positive or negative and is approximately proportional to the reciprocal of their difference. Since the values of the λ 's are not in general known with extreme accuracy, the accuracy of their difference becomes extremely poor as that difference becomes small with respect to the λ 's themselves. Hence in parent-daughter systems in which the two components have comparable half-lives, it is very desirable to count with an instrument which is as specific for the detection of the parent as possible. If this is not practicable, it may be advisable to base the experiment on a direct measurement of the daughter activity, as discussed in the next section.

It should be noted especially that when the daughter is considerably longer lived than the parent ($\lambda_1 > \lambda_2$), the value of g is less than 1, that is, that the presence of a long-lived daughter actually decreases the apparent activity of the parent (see Fig. II.2). It is well known that the presence of a short-lived daughter increases the apparent activity of the parent, since at equilibrium every disintegration of the parent is shortly followed by a disintegration of the daughter, but the opposite effect of a long-lived daughter is likely to be forgotten, since it is not nearly as obvious. Presence of a short-lived daughter will usually cause an actual increase in the counting rate at early times which is quite obviously different from a simple two-component decay; the decay of a system containing a long-lived daughter cannot be distinguished from that of a simple mixture of species.

5. Use of Daughter Activities to Measure Nuclear Reactions

It is sometimes convenient to obtain results from an experiment by analyzing, not for material formed directly, but for a daughter species which may be longer lived, less subject to errors or interferences, or more tractable chemically. In this case the interpretation of the results is straightforward, but requires some thought. We may define two separate cases: case 1, in which the parent is purified and the daughter is allowed to grow into the purified sample, and then is separated and counted; and case 2, in which the parent is not separated, but the daughter is separated directly from the irradiated material.

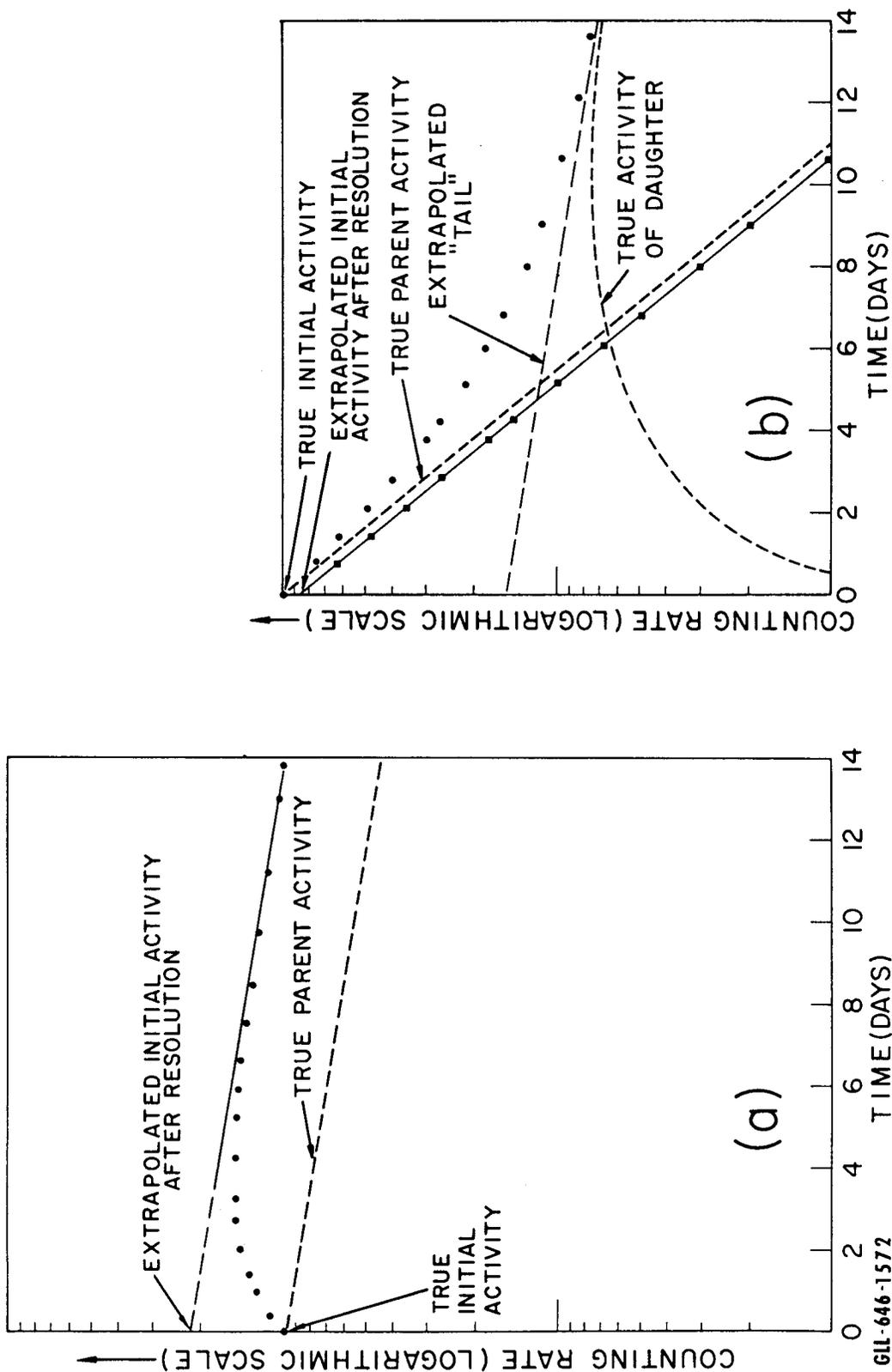


Fig. II.2. Curves showing the early part of the decay of a parent-daughter pair. In each case the half-lives are 12.8 and 1.67 days; in (a) the longer-lived species is the parent, while in (b) the shorter-lived species is the parent. Equal counting efficiencies have been assumed. In neither case does the extrapolated "resolved parent activity" equal the true value. The ratio of observed to true value is given by the factor g (equation II.24). Note especially that, for (b), g is less than 1.

GIL-646-1572

Case 1: This technique is usually used when the half-life of the parent is much longer than that of the daughter. A typical example is the usual method for the analysis of Sr^{90} . In ordinary fission product samples the radiations of Sr^{90} constitute a very small fraction of the total radiations of the strontium fraction. Furthermore, one of the competing species (Sr^{89}) has a half-life of almost two months, so that waiting for the shorter components to decay away, leaving relatively pure Sr^{90} , is out of the question. The analysis may readily be performed, however, by separation and counting of the daughter species Y^{90} , since Sr^{89} has no active daughter. It is well to remark that, prior to purification of the strontium fraction, it is necessary to wait for several days after the irradiation in order to permit the essentially complete decay of Sr^{91} , which is initially present in very large amounts and which does have an active yttrium daughter, as will be discussed under case 2 below.

The analysis is carried out as follows:

1) At the end of the irradiation, the amount of the parent species is (from equation II.10a)

$$x_{10} = \frac{\phi n \sigma}{\lambda_1} \left(1 - e^{-\lambda_1 \Delta} \right) \quad (\text{II.25})$$

where Δ is the duration of the irradiation.

2) After a waiting period of duration w , the parent species is purified. The amount present at this time is (from equations II.25 and II.2)

$$x_{1w} = \frac{\phi n \sigma}{\lambda_1} \left(1 - e^{-\lambda_1 \Delta} \right) e^{-\lambda_1 w} \quad (\text{II.26})$$

3) The purified parent material is allowed to stand for a time t to enable the daughter species to grow into the sample. The time t is usually, though not always, long compared to the half-life of the daughter, since this simplifies the analysis of the results. After the time t has elapsed, the daughter species is purified and counted. The amount present at the time of purification is (from equations II.26 and II.19a)

$$x_2 = \frac{\phi n \sigma}{\lambda_2 - \lambda_1} \left(1 - e^{-\lambda_1 \Delta} \right) e^{-\lambda_1 w} \left[e^{-\lambda_1 t} - e^{-\lambda_2 t} \right] \quad (\text{II.27a})$$

and its counting rate is

$$\text{Counting rate} = \frac{\phi n \sigma \eta_2 \lambda_2}{\lambda_2 - \lambda_1} \left(1 - e^{-\lambda_1 \Delta} \right) e^{-\lambda_1 w} \left(e^{-\lambda_1 t} - e^{-\lambda_2 t} \right) \quad (\text{II.27b})$$

where, as usual, the counting rate is corrected to zero time – in this case, the time of separation from the parent.

In the ideal case, when the parent species is extremely long-lived and the growth period t is long compared to the half-life of the daughter, the expression simplifies to

$$\text{Counting rate} = \frac{\phi n \sigma \eta_2 \lambda_2}{\lambda_2 - \lambda_1} \quad (\text{Ideal Case}) \quad (\text{II.27c})$$

Equation II.27b can be rearranged to give information concerning ϕ , σ , or n in the manner of equations II.10b, II.10c, or II.10e.

Case 2: This technique is usually used when the species directly formed in the reaction of interest is relatively short-lived, decaying into a longer lived species which is more convenient to work with for any one of a number of reasons. A typical example is the usual analysis for Sr^{91} formed in fission. This species is most conveniently measured by analyzing for the daughter product Y^{91} , which is longer lived and less subject to interference from isotopic activities. The analysis is carried out in a different manner from that described under case 1, as follows:

1) During an irradiation of duration Δ , the amount of the parent species builds up according to equation II.10a, so that

$$x_{10} = \frac{\phi n \sigma}{\lambda_1} \left(1 - e^{-\lambda_1 \Delta} \right). \quad (\text{II.28})$$

The daughter species also builds up, in a manner to be derived as follows:

$$\frac{dx_2}{dt} = -\lambda_2 x_2 + \lambda_1 x_1 = -\lambda_2 x_2 + \phi n \sigma \left(1 - e^{-\lambda_1 t} \right). \quad (\text{II.29})$$

Using a method of integration like that used for equation II.6a, letting t go from zero to Δ and letting x_{20} be zero,

$$x_{2\Delta} = \phi n \sigma \left[\frac{1 - e^{-\lambda_2 \Delta}}{\lambda_2} - e^{-\lambda_2 \Delta} \frac{1 - e^{-(\lambda_1 - \lambda_2) \Delta}}{\lambda_1 - \lambda_2} \right]. \quad (\text{II.30})$$

2) The sample is allowed to wait for a time w and the daughter species is then purified and counted. In contrast to case 1, the parent species is never purified. Equation II.19a then applies, with x_{20} of II.19a replaced by $x_{2\Delta}$ from equation II.30, x_{10} obtained from equation II.28, and $t - t_0$ equal to w . This gives, at the purification time,

$$x_2 = \phi n \sigma \left[\frac{1 - e^{-\lambda_2 \Delta}}{\lambda_2} - e^{-\lambda_2 \Delta} \frac{1 - e^{-(\lambda_1 - \lambda_2) \Delta}}{\lambda_1 - \lambda_2} \right] e^{-\lambda_2 w} + \frac{\phi n \sigma}{\lambda_2 - \lambda_1} \left(1 - e^{-\lambda_1 \Delta} \right) \left(e^{-\lambda_1 w} - e^{-\lambda_2 w} \right). \quad (\text{II.31})$$

If the daughter activity is extrapolated back to the end of the irradiation in the usual way, we get

$$x_2 = \phi n \sigma \left\{ \frac{1 - e^{-\lambda_2 \Delta}}{\lambda_2} - e^{-\lambda_2 \Delta} \frac{1 - e^{-(\lambda_1 - \lambda_2) \Delta}}{\lambda_1 - \lambda_2} + \frac{(1 - e^{-\lambda_1 \Delta}) [1 - e^{-(\lambda_1 - \lambda_2) w}]}{\lambda_1 - \lambda_2} \right\} \quad (\text{II.32a})$$

and

$$\text{Counting rate} = \eta_2 \lambda_2 \phi n \sigma \left\{ \frac{1 - e^{-\lambda_2 \Delta}}{\lambda_2} - e^{-\lambda_2 \Delta} \frac{1 - e^{-(\lambda_1 - \lambda_2) \Delta}}{\lambda_1 - \lambda_2} + \frac{(1 - e^{-\lambda_1 \Delta}) [1 - e^{-(\lambda_1 - \lambda_2) w}]}{\lambda_1 - \lambda_2} \right\}. \quad (\text{II.32b})$$

If λ_1 is sufficiently large compared to λ_2 , and w is sufficiently large compared to the parent half-life, then equation II.32b is well approximated by equation II.10a. This demonstrates that formation of a species by way of a negligibly short intermediate species is not appreciably different from direct formation. It should be noted that equations II.32b and II.27b are not equivalent.

An important subcase of case 2 should also be considered. It sometimes happens, for example in the formation of the primary fission products or in reactions where the product formed has a metastable isomeric state, that both the parent and the daughter species are formed directly by nuclear reactions. Analysis of the data in such a case is far from simple. Equation II.29 must be modified to

$$\frac{dx_2}{dt} = -\lambda_2 x_2 + \phi n \sigma_1 \left(1 - e^{-\lambda_1 t}\right) + \phi n \sigma_2 \quad (\text{II.33})$$

which upon integration gives

$$x_2 = \phi n \left[(\sigma_1 + \sigma_2) \frac{1 - e^{-\lambda_2 \Delta}}{\lambda_2} - \sigma_1 e^{-\lambda_2 \Delta} \frac{1 - e^{-(\lambda_1 - \lambda_2) \Delta}}{\lambda_1 - \lambda_2} \right]. \quad (\text{II.34})$$

Treating the sample in the same manner as before, we arrive at the results

$$x_2 = \phi n \left\{ (\sigma_1 + \sigma_2) \frac{1 - e^{-\lambda_2 \Delta}}{\lambda_2} - \sigma_1 e^{-\lambda_2 \Delta} \frac{1 - e^{-(\lambda_1 - \lambda_2) \Delta}}{\lambda_1 - \lambda_2} + \sigma_1 \frac{(1 - e^{-\lambda_1 \Delta}) \left[1 - e^{-(\lambda_1 - \lambda_2) w}\right]}{\lambda_1 - \lambda_2} \right\}, \quad (\text{II.35a})$$

$$\text{Counting rate} = \eta_2 \lambda_2 \phi n \left\{ (\sigma_1 + \sigma_2) \frac{1 - e^{-\lambda_2 \Delta}}{\lambda_2} - \sigma_1 e^{-\lambda_2 \Delta} \frac{1 - e^{-(\lambda_1 - \lambda_2) \Delta}}{\lambda_1 - \lambda_2} + \sigma_1 \frac{(1 - e^{-\lambda_1 \Delta}) \left[1 - e^{-(\lambda_1 - \lambda_2) w}\right]}{\lambda_1 - \lambda_2} \right\}. \quad (\text{II.35b})$$

We then see that if the experiment is intended to give a measurement of cross section, two runs must be made with times Δ and w selected to give substantially different values of the exponential factor, since two separate cross sections are involved; furthermore, we see that unless the later terms in the bracket are negligibly small compared to the first term (which, for comparable cross sections, will only happen if $\lambda_1 \gg \lambda_2$, that is, for a very short-lived parent), the effective reaction cross section is not simply the sum of the two cross sections involved, but depends on the time schedule of the experiment in a rather elaborate fashion. In cases where the half-lives are comparable to each other and to w and Δ , the errors caused by this effect may be large indeed.

If the two species concerned are nuclear isomers, purification of the daughter from the parent is not usually practical in an analytical procedure. In this case equation II.35b still applies, with the value of w taken as infinite.

It is also of interest to note that if both species can be counted, either in the same sample or separately, then the counting rate of the parent species furnishes a measure of σ_1 if equations II.24 and II.10b are used. This value may then be used in equation II.35b to permit the value of σ_2 to be calculated.

6. Branching Decay

Some radioactive species may follow either of two decay paths to produce either of two decay products. (Am^{242} , indeed, follows three!) The ratio between the rates of the two kinds of disintegrations is fixed and characteristic of the species, and is known as the branching ratio. Many examples are known in the natural radioactive series - UX_1 (Th^{234}) may decay by beta emission to either one of the two isomers UX_2 ($\text{Pa}^{231\text{m}}$) or UZ (Pa^{231}); or RaC (Bi^{214}) may decay by beta emission to RaC' (Po^{214}) or by alpha emission to RaC'' (Tl^{210}). This behavior may be described mathematically by separating the decay constant λ of equation II.1 into two partial decay constants, λ_{1a} and λ_{1b} , whose ratio is the branching ratio.

There is no effect of branching decay on the time dependence of the counting rate if neither of the daughter species is active, although the counter efficiency is no longer simple:

$$\eta_1 = \frac{\lambda_{1a}}{\lambda_{1a} + \lambda_{1b}} \eta_{1a} + \frac{\lambda_{1b}}{\lambda_{1a} + \lambda_{1b}} \eta_{1b}. \quad (\text{II.36})$$

If the daughter species are active, equations II.1 and II.18 become

$$\frac{dx_1}{dt} = -(\lambda_{1a} + \lambda_{1b}) x_1, \quad (\text{II.37})$$

$$\frac{dx_{2a}}{dt} = -\lambda_{2a} x_{2a} + \lambda_{1a} x_1, \quad (\text{II.38a})$$

$$\frac{dx_{2b}}{dt} = -\lambda_{2b} x_{2b} + \lambda_{1b} x_1. \quad (\text{II.38b})$$

Upon integration, these equations give

$$x_1 = x_{10} e^{-(\lambda_{1a} + \lambda_{1b})(t-t_0)}, \quad (\text{II.39a})$$

$$x_{2a} = x_{2a0} e^{-\lambda_{2a}(t-t_0)} + \frac{\lambda_{1a} x_{10}}{\lambda_{2a} - \lambda_{1a} - \lambda_{1b}} \left[e^{-(\lambda_{1a} + \lambda_{1b})(t-t_0)} - e^{-\lambda_{2a}(t-t_0)} \right], \quad (\text{II.39b})$$

$$x_{2b} = x_{2b0} e^{-\lambda_{2b}(t-t_0)} + \frac{\lambda_{2b} x_{10}}{\lambda_{2b} - \lambda_{1a} - \lambda_{1b}} \left[e^{-(\lambda_{1a} + \lambda_{1b})(t-t_0)} - e^{-\lambda_{2b}(t-t_0)} \right], \quad (\text{II.39c})$$

and therefore

$$\begin{aligned} \text{Counting rate} &= \eta_1 (\lambda_{1a} + \lambda_{1b}) x_{10} e^{-(\lambda_{1a} + \lambda_{1b})(t-t_0)} + \eta_{2a} \lambda_{2a} x_{2a0} \\ &\times e^{-\lambda_{2a}(t-t_0)} + \eta_{2b} x_{2b0} \lambda_{2b} e^{-\lambda_{2b}(t-t_0)} + \frac{\lambda_{1a} \lambda_{2a} \eta_{2a} x_{10}}{\lambda_{2a} - \lambda_{1a} - \lambda_{1b}} \\ &\times \left[e^{-(\lambda_{1a} + \lambda_{1b})(t-t_0)} - e^{-\lambda_{2a}(t-t_0)} \right] + \frac{\lambda_{1b} \lambda_{2b} \eta_{2b} x_{10}}{\lambda_{2b} - \lambda_{1a} - \lambda_{1b}} \left[e^{-(\lambda_{1a} + \lambda_{1b})(t-t_0)} \right. \\ &\quad \left. - e^{-\lambda_{2b}(t-t_0)} \right]. \end{aligned} \quad (\text{II.40})$$

The apparent counting rate of the parent species contains a growth factor g which is calculated as before by dividing the observed counting rate of the component with the parent half-life by the counting rate calculated for the pure parent species, to give

$$\begin{aligned} g &= 1 + \left(\frac{\lambda_{1a}}{\lambda_{1a} + \lambda_{1b}} \right) \left(\frac{\lambda_{2a}}{\lambda_{2a} - \lambda_{1a} - \lambda_{1b}} \right) \frac{\eta_{2a}}{\eta_1} + \left(\frac{\lambda_{1b}}{\lambda_{1a} + \lambda_{1b}} \right) \\ &\quad \times \left(\frac{\lambda_{2b}}{\lambda_{2b} - \lambda_{1a} - \lambda_{1b}} \right) \frac{\eta_{2b}}{\eta_1} \end{aligned} \quad (\text{II.41})$$

which is analogous to equation II.24. If only one daughter species is active, η_{2b} may be set equal to zero in all the above equations.

Branching decay systems are frequently described in terms of partial half-lives, defined by

$$T_{\frac{1}{2}a} = \frac{\ln 2}{\lambda_{1a}} = \frac{\lambda_{1a} + \lambda_{1b}}{\lambda_{1a}} T_{\frac{1}{2}} \quad (\text{II.42a})$$

$$T_{\frac{1}{2}b} = \frac{\ln 2}{\lambda_{1b}} = \frac{\lambda_{1a} + \lambda_{1b}}{\lambda_{1b}} T_{\frac{1}{2}} \quad (\text{II.42b})$$

7. Decay Chains With More Than Two Consecutive Decays

The mathematical techniques applied to two-membered decay chains can readily be extended to multimembered chains, including independent formation and branched decay if necessary. The possible permutations and combinations are so numerous that it is not practical to investigate them in detail; it is enough to recapitulate the mathematical technique used, and to describe another technique which may simplify the labor of deriving the sometimes very elaborate formulae which describe the behavior of such complex systems.

The method of derivation is summarized:

1) Write the differential equation for each member of the chain, as follows: The first time derivative of the amount x_i of each species i is equal to the sum of the following three terms:

- a) The negative of the product of x_i by the total decay constant λ_i of species i .
- b) The rate of formation of the species i by any induced nuclear reactions occurring in the system. This expression will in general be the product of a reaction cross section σ_i , a number of atoms n_i of target material, and a flux ϕ_i . If more than one reaction is going on, the term for each reaction is written and all such terms are summed. ϕ_i and n_i may be functions of time - for example, the target material may be a species which is itself formed by a nuclear reaction. If x_i is being consumed in a nuclear reaction, the term for this reaction is given a negative sign and x_i is used in place of n_i .
- c) The rate of formation of the species i by radioactive decay from one or more parent species. These terms are always given by the product of the number of atoms of the parent

species present (which is a function of time) by the partial decay constant of that parent species for formation of the species i . Terms from various parents are additive.

2) Start with the equation for the first member of the decay chain. This equation can always be integrated directly, like equation II.1, or by the use of an integrating factor, like equation II.6a, depending on the form of the equation.

3) Substitute the expression obtained into the equations for each member of the second generation.

4) Rearrange the terms in the equations so obtained as in equation II.6b and integrate in the same manner to obtain expressions for the amount of each second-generation species present.

5) Substitute these expressions into the equations for the third-generation species and so on.

6) Multiply the amount of each species by its decay constant and counting efficiency to obtain the counting rate of each species as a function of time. Sum the counting rates where applicable to get the total counting rate of the sample.

7) For each change in the experimental conditions — stopping of an irradiation, purification of a parent species followed by subsequent counting, and the like — use the above technique to give the number of atoms of each species present at the instant of the change; then use this expression as the starting condition for that species and repeat the derivation, using suitably modified differential equations.

The integrands to be integrated will never be more elaborate than simple exponentials unless the system is exposed to a time-dependent flux.

A second approach, which involves somewhat less labor for complex chains, is based on the fact that the amount of any species in a system is always expressible by a constant term (which may be zero) plus a sum of exponential terms involving only the total decay constants of that species and its progenitors in the system, and no others. These decay constants never appear in combinations in the exponents. We can therefore proceed as follows:

1) Write the differential equations of the system, as before.

2) Express the amount of each species present as a linear combination of the proper exponentials with undetermined coefficients, plus an undetermined constant term.

3) Apply the known starting conditions of the system to give a set of equations relating the unknown coefficients.

4) Substitute the assumed expressions in the differential equations.

5) Obtain additional relations between the coefficients by setting the coefficient of each exponential term in the resulting equations, and the constant term, equal to zero. These relations, together with those obtained in step 3 above, will always be enough to determine the coefficients.

6) Calculate counting rates as before.

An example follows. Let us assume that species 1 is formed at constant rate in a nuclear reaction; that it decays by branching decay to two species 2a and 2b; that each of these decays to the same active species 3, which decays to a stable product; and that the starting condition is that the original system contained no active material. For simplicity, we shall measure time from the start of the reaction.

The differential equations are:

$$\frac{dx_1}{dt} = -(\lambda_{1a} + \lambda_{1b}) x_1 + \phi n \sigma, \quad (\text{II.43a})$$

$$\frac{dx_{2a}}{dt} = -\lambda_{2a} x_{2a} + \lambda_{1a} x_1, \quad (\text{II.43b})$$

$$\frac{dx_{2b}}{dt} = -\lambda_{2b} x_{2b} + \lambda_{1b} x_1, \quad (\text{II.43c})$$

$$\frac{dx_3}{dt} = -\lambda_3 x_3 + \lambda_{2a} x_{2a} + \lambda_{2b} x_{2b}. \quad (\text{II.43d})$$

We assume that

$$x_1 = A_1 + B_1 e^{-(\lambda_{1a} + \lambda_{1b})t}, \quad (\text{II.44a})$$

$$x_{2a} = A_{2a} + B_{2a} e^{-(\lambda_{1a} + \lambda_{1b})t} + C_{2a} e^{-\lambda_{2a}t}, \quad (\text{II.44b})$$

$$x_{2b} = A_{2b} + B_{2b} e^{-(\lambda_{1a} + \lambda_{1b})t} + C_{2b} e^{-\lambda_{2b}t}, \quad (\text{II.44c})$$

$$x_3 = A_3 + B_3 e^{-(\lambda_{1a} + \lambda_{1b})t} + C_{3a} e^{-\lambda_{2a}t} + C_{3b} e^{-\lambda_{2b}t} + D_3 e^{-\lambda_3 t}, \quad (\text{II.44d})$$

and therefore from the starting conditions ($t = 0$)

$$A_1 + B_1 = 0, \quad (\text{II.45a})$$

$$A_{2a} + B_{2a} + C_{2a} = 0, \quad (\text{II.45b})$$

$$A_{2b} + B_{2b} + C_{2b} = 0, \quad (\text{II.45c})$$

$$A_3 + B_3 + C_{3a} + C_{3b} + D_3 = 0. \quad (\text{II.45d})$$

Substituting the values of the x 's from equations II.44a through II.44d into equations II.43a through II.43d, performing the indicated differentiations, and collecting terms, we get

$$0 = -(\lambda_{1a} + \lambda_{1b})A_1 + \phi n \sigma, \quad (\text{II.46a})$$

$$0 = (\lambda_{1a} A_1 - \lambda_{2a} A_{2a}) + \left[\lambda_{1a} B_1 - (\lambda_{2a} - \lambda_{1a} - \lambda_{1b}) B_{2a} \right] e^{-(\lambda_{1a} + \lambda_{1b})t}, \quad (\text{II.46b})$$

$$0 = (\lambda_{1b} A_1 - \lambda_{2b} A_{2b}) + \left[\lambda_{1b} B_1 - (\lambda_{2b} - \lambda_{1a} - \lambda_{1b}) B_{2b} \right] e^{-(\lambda_{1a} + \lambda_{1b})t}, \quad (\text{II.46c})$$

$$0 = (\lambda_{2a} A_{2a} + \lambda_{2b} A_{2b} - \lambda_3 A_3) + \left[(\lambda_{1a} + \lambda_{1b} - \lambda_3) B_3 + \lambda_{2a} B_{2a} + \lambda_{2b} B_{2b} \right] \\ \times e^{-(\lambda_{1a} + \lambda_{1b})t} + \left[(\lambda_{2a} - \lambda_3) C_{3a} + \lambda_{2a} C_{2a} \right] e^{-\lambda_{2a}t} \\ + \left[(\lambda_{2b} - \lambda_3) C_{3b} + \lambda_{2b} C_{2b} \right] e^{-\lambda_{2b}t}. \quad (\text{II.46d})$$

From equation II.46a we get A_1 and from that and equation II.45a we get B_1 . We then let the coefficients of each exponential term, and the constant term, equal zero separately in equations II.46b and II.46c to give

$$\lambda_{1a} A_1 - \lambda_{2a} A_{2a} = 0, \quad (\text{II.47a})$$

$$\lambda_{1a} B_1 - (\lambda_{2a} - \lambda_{1a} - \lambda_{1b}) B_{2a} = 0, \quad (\text{II.47b})$$

$$\lambda_{1b} A_1 - \lambda_{2b} A_{2a} = 0, \quad (\text{II.47c})$$

$$\lambda_{1b} B_1 - (\lambda_{2b} - \lambda_{1a} - \lambda_{1b}) B_{2b} = 0. \quad (\text{II.47d})$$

Substituting the determined expressions for A_1 and B_1 we get expressions for A_{2a} , B_{2a} , A_{2b} , and B_{2b} , and by substituting these in equations II.45b and II.45c we get expressions for C_{2a} and C_{2b} . We then let the coefficient of each exponential term in equation II.46d, and the constant term, each equal zero and by substituting in the expressions already determined for the first six coefficients we get expressions for all the rest except D_3 , which is then obtained from equation II.45d. Some labor may often be saved by evaluating only the minimum number of coefficients required to give the desired information — that is to say, the factors making up the observed counting rate of the components being analyzed in the experiment.

It should be emphasized that in any system, whether a simple mixture of species or a complicated tangle of consecutive or branching decays, the decay of the activity of the system can be described by a sum of exponential terms, one component for each species in the system; and that the half-lives of the components are identical with the half-lives of the species in the system. Knowledge of the growth and decay formulae derived in this chapter is not necessary for the resolution of such data into components; it is, however, most definitely necessary for the interpretation of the resolved data in terms of measurement of fluxes, cross sections, or other physical properties associated with the experiment.

III. MATHEMATICAL TECHNIQUES

1. Nature of the Problem

Measurements of radioactivity are, like all physical measurements, subject to both random and systematic error. Unlike many physical measurements, however, measurements of radioactivity are plagued with inherent, unavoidable random errors of quite appreciable size which are due to the very nature of the phenomenon and cannot be eliminated by any refinement of technique. As a result, the elaborate formulae of the preceding chapter must be considered to represent an unattainable ideal. Our initial assumption of precise data must be abandoned.

We are left with the rather thorny problem of developing an objective technique for using admittedly imperfect data to make a "best" estimate of the quantities of interest in the formulae of Chapter II; of developing a definition of "best" in this sense which is both objective and justifiable; and, further, of forming an estimate of the reliability of the estimates so obtained. In this chapter it has been my objective to collect in one place an assortment of mathematical techniques which apply to the problem. Full development of these disciplines has not been attempted and is, indeed, beyond my competence; I have endeavored to include only those facets of the subject which apply more or less directly to the analysis of data on radioactivity.

2. Matrix Algebra

Considerable use will be made of matrix algebra and matrix-vector notation in our subsequent development, since the notation is both concise and elegant.

A matrix is simply a rectangular ordered array of numbers arranged in horizontal rows and vertical columns. The coefficients of a system of simultaneous linear equations, for example, constitute a matrix. If we regard a collection of matrices as a set of mathematical entities, and wish to develop an algebra to handle relationships between them, it is necessary to postulate a set of self-consistent axioms. An algebra can be developed on the basis of any such set; since it is extremely useful to relate the algebra of matrices to the more ordinary algebra of linear equations, our axioms will be derived from those of ordinary algebra through consideration of sets of linear equations.

Let us consider a set of simultaneous linear equations. Such equations consist of statements that the value of a specified linear combination of the members of a set of quantities x_i is equal to one of the members of another set of quantities, y_j . If all the x_i are known, an unlimited number of y_j can be derived by varying the coefficients of the x_i in the linear combinations.

Let us consider a set of m linear equations of the form

$$\sum_{i=1}^n a_{ji} x_i = y_j, \quad 1 \leq j \leq m. \quad (\text{III.1})$$

The set of coefficients a_{ji} may be represented by the matrix A, where the quantity a_{ji} is the number in the j th row and i th column of A. A convenient notation for an element of the matrix A is A_{ji} , the subscripts signifying the row and column, in that order, in which the element is located. We have then for a description of the elements of the matrix A

$$A_{ji} = a_{ji}. \quad (\text{III.2})$$

We then wish to express the set of x_i and the set of y_j in matrix form, and define rules of combination of matrices so that the very compact equation

$$AX = Y \quad (\text{III.3})$$

is a valid matrix expression equivalent to the complete set of equations III.1 for all values of the subscript j . We have a choice between expressing the sets of x_i and y_j as matrices of one row and expressing them as matrices of one column; the choice is arbitrary. In accordance with convention, we shall describe both the set x_j and the set y_j as matrices of one column; X has as many elements as there are columns of A, and Y has as many elements as there are rows of A. A matrix of one column is usually referred to as a vector; we shall define the operation of matrix-vector multiplication as indicated in equation III.3 by

$$\sum_{i=1}^n A_{ji} X_i = Y_j \quad (\text{III.4})$$

which is obviously identical with equation III.1, except that the quantities concerned are described by their locations in the appropriate matrices rather than by their magnitudes. The operation is not defined unless the number of elements of X is the same as the number of columns of A. In particular, the operation XX is not defined.

The matrix formed from the matrix A by interchanging rows and columns is called the transpose of A, symbolized by A^t . In other words,

$$A_{ij}^t = A_{ji}. \quad (\text{III.5})$$

The transpose of a vector is a matrix of one row (sometimes called a row vector as opposed to a column vector) and the operation $X^t X$ is defined; in fact

$$X^t X = \sum_{j=1}^n X_j^2 \quad (\text{III.6})$$

from equation III.4. If X is considered an actual vector in Euclidean n -space with components of X_j , the quantity $X^t X$ is the square of the amplitude of that vector. The operation XX^t is also defined; the result is a square matrix whose ij component is $X_i X_j$.

If in addition to the relationship given by equation III.3 we have a relationship

$$BY = Z, \quad (\text{III.7})$$

then obviously

$$Z = BAX. \quad (\text{III.8})$$

We can also derive a relationship between Z and X directly of the form

$$Z = CX \quad (\text{III.9})$$

so that for consistency it is desirable to define the operation of multiplication of the matrices B and A in such a manner that

$$C = BA \quad (\text{III.10})$$

where the sign of equality implies that each element of the matrix formed by multiplication of the matrices B and A is equal to the corresponding element of the matrix C .

It is a simple matter to show that the consistent definition of the operation described in equation III.10 is, for a matrix B with n columns and a matrix A with n rows,

$$C_{jk} = \sum_{i=1}^n B_{ji} A_{ik}. \quad (\text{III.11})$$

The number of columns of B must equal the number of rows of A for the operation to be defined. It can easily be seen that matrix multiplication as defined by equation III.11 is equivalent to matrix-vector multiplication as defined by equation III.1, considering the columns of the matrix A as a set of vectors each of which is transformed in turn into the corresponding column of the matrix C by multiplication by the matrix B . Matrix multiplication is not commutative, that is, AB does not necessarily equal BA , but can be shown to be associative, i. e. $(AB)C = A(BC)$.

Some special forms of matrices are of interest: A square matrix is a matrix with equal numbers of rows and columns; a diagonal matrix is a square matrix whose ij elements are all zero unless $i = j$; an upper triangular matrix is a square matrix whose ij elements are all zero unless $j \geq i$; a lower triangular matrix is a square matrix whose ij elements are all zero unless $i > j$; a zero matrix O is a matrix whose elements are all zero; a unit matrix I is a diagonal matrix whose nonzero elements are all unity; and a symmetric matrix is a matrix such that $A_{ij} = A_{ji}$, hence $A = A^t$.

The determinant of a square matrix is the determinant whose elements are the elements of that matrix. Its value is, according to the theory of determinants,

$$|A| = \sum_{j, k}^{m!} \left[(-1)^n \prod_{j, k}^m A_{jk} \right] \quad (\text{III.12})$$

where within each product of m factors j and k each take on each of the values 1 to m once and only once, the values of j are taken in order, the sum is taken over all sets of permutations of k , and the exponent n is the number of interchanges of the subscripts k necessary to restore them to normal order. If the value of the determinant of a square matrix is zero, the matrix is said to be singular.

It can be shown that a nonsingular square matrix A possesses an inverse A^{-1} such that the product of the two, in either order, is equal to a unit matrix of the same dimensions. In order to evaluate the inverse of a given matrix, we define the cofactor of an element of the matrix A_{ij} to be the determinant of the matrix formed by eliminating the row i and column j , multiplied by $(-1)^{i+j}$. The cofactor of A_{ij} is symbolized by A^{ij} . It will be recalled, from the theory of determinants, that

$$|A| = \sum_{i=1}^m A_{ij} A^{ij} = \sum_{j=1}^m A_{ij} A^{ij}. \quad (\text{III.13})$$

In the theory of determinants this is usually spoken of as expansion by minors. From equation III.13 and the definition of matrix multiplication (equation III.11) it is easy to see that

$$(A^{-1})_{ij} = \frac{A^{ji}}{|A|} \quad (\text{III.14})$$

or, that the inverse matrix is formed by replacing each element by the value of its cofactor, dividing each element by the determinant of the entire matrix, and then (note the reversal of indices) transposing the resulting matrix.

From these basic definitions a number of simple relationships become apparent, as follows:

$$(\alpha A)_{ij} = \alpha A_{ij} \quad (\text{III.15})$$

where α is a simple scalar quantity;

$$(A + B)_{ij} = A_{ij} + B_{ij}, \quad (\text{III.16})$$

$$A + B = B + A, \quad (\text{III.17})$$

$$A(BC) = (AB)C = ABC, \quad (\text{III.18})$$

$$A(B + C) = AB + AC, \quad (\text{III.19})$$

$$(A + B) + C = A + (B + C) = A + B + C. \quad (\text{III.20})$$

If A is any matrix, $A^t A$ is a square symmetric matrix with as many columns as A , and AA^t is a square symmetric matrix with as many rows as A . (III.21)

$$AI = IA = A, \quad (\text{III.22})$$

$$A \cdot O = O \cdot A = O, \quad (\text{III.23})$$

$$(AB)^t = B^t A^t, \quad (\text{III.24})$$

$$(AB)^{-1} = B^{-1} A^{-1}. \quad (\text{III.25})$$

We define the term orthogonal as follows: if two vectors X and Y have the same number of components, they are orthogonal if

$$X^t Y = Y^t X = 0. \quad (\text{III.26})$$

Returning to our set of linear equations, as expressed in equation III.3, if there are m equations of m terms each, and the m components of Y are known while the m components of X are not, we can in principle evaluate the components of X by multiplying both sides of the equation through from the left by A^{-1} , if A is not singular, to give

$$A^{-1} Y = A^{-1} AX = IX = X. \quad (\text{III.27})$$

We shall use this technique in our data analysis; as will be shown shortly, the values of the elements of the inverse matrix will be extremely important to our problem.

3. Statistical Mathematics

We make the assumption that the results of an experiment subject to random error, if repeated sufficiently often, will "average out" to the "true" result of an ideal experiment not subject to random errors. Intuitively we feel that most of our actual results will be close to the ideal value, fewer of them somewhat less close, and a very small number quite far away indeed. This leads us to the concept of a distribution function for an experimental measurement subject to random errors; such a measurement will be referred to as a variate. The distribution function for a variate x is a function which describes the fraction of times, in a large number of measurements of x , that a measurement of x falls between two specified values. "Fraction of times in a large number of measurements" may be taken as an approximate definition of probability - the rigorous definition may be taken as the limit of that fraction as the number of measurements increases without limit. The utility of the concept stems from the fact that probabilities may often be calculated from first principles.

Distributions can of course arise from other sources than from errors in the measurement of some simple precise quantity. If one measures the weights of a number of supposedly identical objects, some variation will be observed; part of this variation may be due to random errors in the weighing process and part to actual differences between the objects themselves. In this case the "averaging out" process carried out over a number of measurements refers, not to the "true" weight of any actual object, but to a sort of idealized or typical weight which one would expect such an object to have, a priori.

It should be emphasized that the concept of "average" applies to a distribution; it makes no sense to talk of the "average" of a set of quantities from different distributions unless some correction is applied to reduce these distributions to the same basis. This concept will be discussed in greater detail below.

The probability of a measurement of the variate x falling in a particular interval is obviously dependent on the size of that interval; as the size of the interval approaches zero, the probability of a measurement falling in that interval becomes simply proportional to the size of the interval.

Let us assume a coordinate u measured in units of a particular dimensionality — area, for example, or time or pure number. We then define the distribution function $P_x(u)$ of the variate x , when x is measured in the same units as u , to be such that $P_x(u) du$ is the probability that a measured value of x will fall between u and $u + du$. We will assume that distribution functions are normalized, that is

$$\int_{-\infty}^{\infty} P_x(u) du = 1. \quad (\text{III.28})$$

This is equivalent to saying that any measurement of x will result in some finite value.

From the definition, the probability that a measured value of x will fall between a and b is

$$P(a \leq x \leq b) = \int_a^b P_x(u) du. \quad (\text{III.29})$$

We define the expected value $E[f(x)]$ of any function of x as the limit of the average value from a large number n of measurements, as n increases without limit; therefore

$$E(x) = \int_{-\infty}^{\infty} u P_x(u) du \quad (\text{III.30})$$

and

$$E[f(x)] = \int_{-\infty}^{\infty} f(u) P_x(u) du. \quad (\text{III.31})$$

If $f(x)$ is taken as the difference between x and $E(x)$, we get

$$E[x - E(x)] = \int_{-\infty}^{\infty} [u - E(x)] P_x(u) du = E(x) - E(x) = 0 \quad (\text{III.32})$$

as might be expected; if, however, we consider the expected value of the square of that difference, we are integrating a quantity which is sometimes

positive and never negative so that we expect a nonzero result. This quantity is called the variance of x :

$$\text{var}(x) = \int_{-\infty}^{\infty} [u - E(x)]^2 P_x(u) du = E \left\{ [x - E(x)]^2 \right\} = \sigma_x^2. \quad (\text{III.33})$$

Its square root is the standard deviation of x , σ_x .

It should be noted that the term "variance of x " is somewhat misleading; $\text{var}(x)$ is not a function of x , but is a parameter of the distribution function of x along the coordinate u . It is a measure of the precision with which x can be measured, or of the "width" of $P_x(u)$. Since the dimensions of $\text{var}(x)$ are those of x^2 , while the dimensions of σ_x are those of x , the standard deviation is the quantity which is the easier to interpret in terms of the precision of x . x is often reported as $x \pm \sigma_x$. If the form of the distribution function is known, $\pm \sigma_x$ gives information on the size of the interval within which x is expected to fall; obviously the size of this interval depends on the degree of certainty with which we wish to predict the value of x , or the result of a measurement to be made. It is customary to speak of a confidence limit to a certain level; the 95% confidence limits, for example, define the range in x within which 95% of a large number of measurements are expected to fall. (See Fig. III.1.)

The "n percent confidence interval" can be defined explicitly by a such that

$$\int_{\bar{x}-a}^{\bar{x}+a} P_x(u) du = \frac{n}{100} \quad (\text{III.33a})$$

which obviously depends on the form of $P_x(u)$. It is the confidence interval which determines the significance of a set of results – for example, the question may be asked whether the results of two measurements are believed to be the same with a certain arbitrarily assigned confidence. This is equivalent to considering the difference between the two measurements to be a new variate whose distribution function can be derived, and asking the question whether the value zero lies within the appropriate confidence interval.

If we have two variates x and y , we define an analogous quantity, the covariance of x and y , by

$$\text{cov}(x, y) = E \{ [x - E(x)][y - E(y)] \}. \quad (\text{III 34})$$

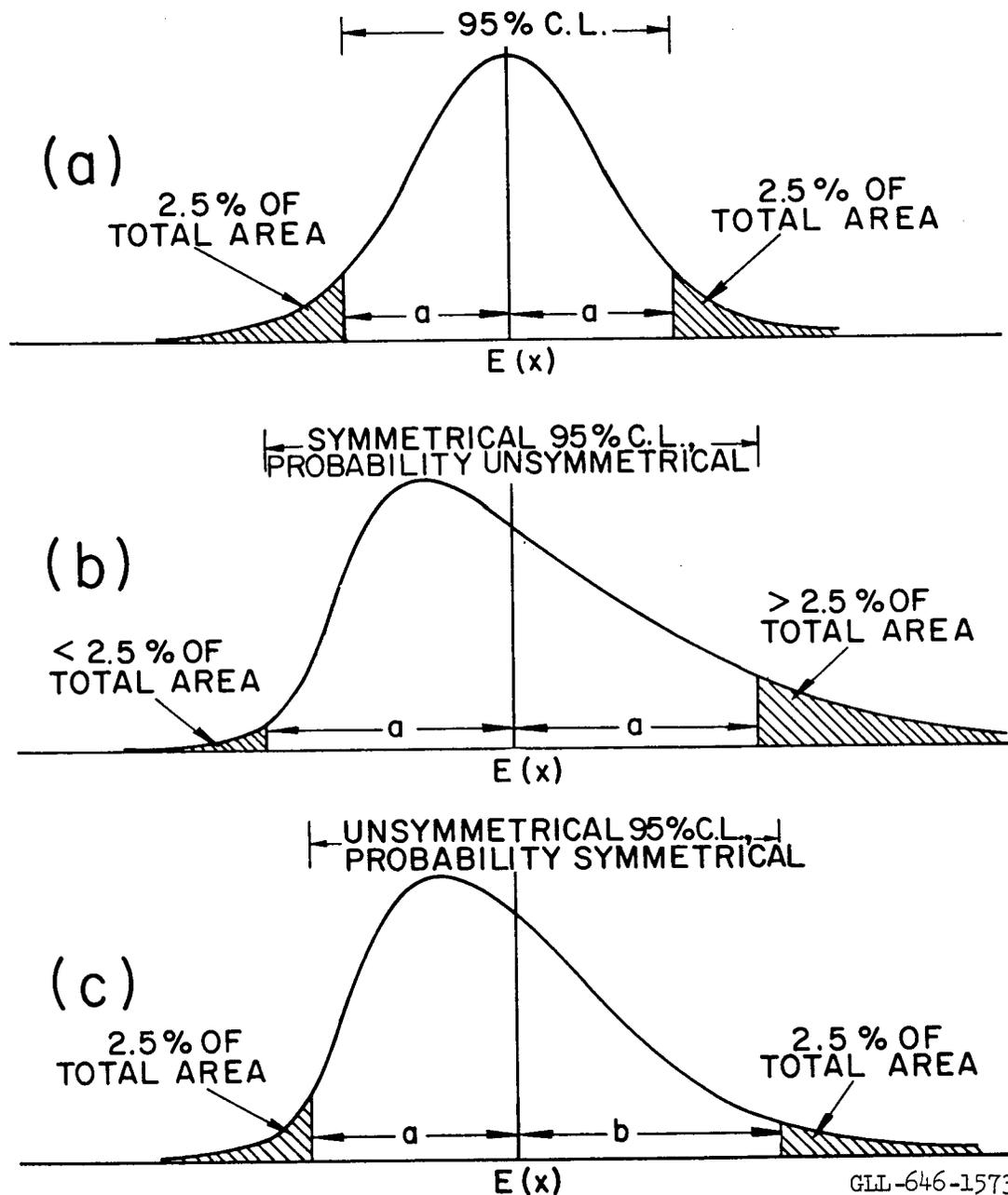


Fig. III.1. Confidence limits: the 95% confidence limit is shown for three distributions. For a symmetrical distribution (a) a symmetrical interval about $E(x)$ excludes 5% of the total probability, in two equal parts. This implies that there is a 95% probability that a single measurement of x will lie between $E(x) + a$ and $E(x) - a$, with an equal probability (of 2.5%) that a measurement will lie outside this range on either side. For an asymmetrical distribution one has a choice: (b) shows a symmetrical interval about $E(x)$ such that the probability of a measurement of x falling within this interval is 95%, but a measurement outside this interval is more likely to be too high than too low; (c) shows an asymmetrically defined interval, so placed that a measurement falling outside this interval is as likely to be high as it is to be low.

Naturally,

$$\text{var}(x) = \text{cov}(x, x). \quad (\text{III.35})$$

If $\text{cov}(x, y) = 0$, x and y are said to be uncorrelated. $\text{Cov}(x, y)$ is a measure of the correlation between deviations from the expected values of x and y , measured simultaneously. If for example one measures both the length and the temperature of each of a set of metal bars, one might not be surprised to find that there was a distribution of both length and temperature; if evaluation of the covariance between length and temperature, however, gives a result which is markedly different from zero, one might then begin to suspect a relationship between the temperature of a bar of metal and its length. One could then get a more realistic idea of the actual distribution of lengths among the metal rods by measuring them all at some constant temperature.

Such a "distribution of lengths" would of course be compounded of the actual distribution in the lengths of the bars and the imprecision of the method of measurement used. If one bar were to be measured a number of times, a distribution would still be obtained, though with a method of measurement sufficiently precise in terms of the unit of measurement such a distribution would be extremely narrow. Information on the precision of the measurement method, obtained by a series of measurements on a single object, can then be used to deduce the actual spread of the distribution in length of a set of bars from the observed total spread. (This deduction will be discussed further under "Analysis of Replicate Samples: Analysis of Variance," Section 7 of Chapter V.)

A number which is not subject to variation is referred to as a constant. It can easily be seen that, for a constant a ,

$$E(a) = a, \quad (\text{III.36})$$

$$E(ax) = aE(x), \quad (\text{III.37})$$

$$E(a + x) = a + E(x), \quad (\text{III.38})$$

$$\text{var}(a) = 0, \quad (\text{III.39})$$

$$\text{cov}(a, x) = \text{cov}(x, a) = 0, \quad (\text{III.40})$$

$$\text{var}(ax) = a^2 \text{var}(x), \quad (\text{III.41})$$

$$\text{var}(a + x) = \text{var } x, \quad (\text{III.42})$$

$$\text{cov}(ax, by) = ab \text{ cov}(x, y). \quad (\text{III.43})$$

$E(x)$, $\text{var}(x)$, and $\text{cov}(x, y)$ are constants of the system.

We are now in a position to develop what we will need of the algebra of expected values, variances, and covariances.

Theorem I: $E(x + y) = E(x) + E(y)$.

$$E(x + y) = \int_{-\infty}^{\infty} u P_{x+y}(u) du,$$

$$P_{x+y}(u) = \int_{-\infty}^{\infty} P_y(v) P_x(u - v) dv,$$

$$E(x + y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u P_y(v) P_x(u - v) dv du.$$

Reversing the order of integration, and letting $u = u - v + v$,

$$E(x + y) = \int_{-\infty}^{\infty} P_y(v) \left[\int_{-\infty}^{\infty} (u - v) P_x(u - v) du + v \int_{-\infty}^{\infty} P_x(u - v) du \right] dv.$$

The inner integrals are being taken over the entire range of u with v held constant so that $du = d(u - v)$ for each value of v .

$$\begin{aligned} \therefore E(x + y) &= \int_{-\infty}^{\infty} P_y(v) \left[\int_{-\infty}^{\infty} (u - v) P_x(u - v) d(u - v) \right. \\ &\quad \left. + v \int_{-\infty}^{\infty} P_x(u - v) d(u - v) \right] dv \end{aligned}$$

$$= \int_{-\infty}^{\infty} P_y(v) [E(x) + v] dv$$

$$= E(x) \int_{-\infty}^{\infty} P_y(v) dv + \int_{-\infty}^{\infty} v P_y(v) dv.$$

$$E(x + y) = E(x) + E(y). \quad \text{Q. E. D.}$$

Theorem II: $E(xy) = E(x)E(y) + \text{cov}(x, y)$.

$$\begin{aligned} \text{cov}(x, y) &= E \left\{ [x - E(x)][y - E(y)] \right\} \\ &= E[xy - xE(y) - yE(x) + E(x)E(y)] \\ &= E(xy) - E(y)E(x) - E(x)E(y) + E(x)E(y) \\ &= E(xy) - E(x)E(y). \end{aligned}$$

$\therefore E(xy) = E(x)E(y) + \text{cov}(x, y)$. Q. E. D.

Corollary: $E(x^2) = \overline{E(x)}^2 + \text{var}(x)$

$$\begin{aligned} E(x^2) &= E(x \cdot x) = E(x)E(x) + \text{cov}(x, x) \\ \therefore E(x^2) &= \overline{E(x)}^2 + \text{var}(x). \quad \text{Q. E. D.} \end{aligned}$$

• Theorem III: $\text{var}(x + y) = \text{var}(x) + \text{var}(y) + 2 \text{cov}(x, y)$.

$$\begin{aligned} \text{var}(x + y) &= E \left\{ [x + y - E(x + y)]^2 \right\} \\ &= E \left\{ [x - E(x) + y - E(y)]^2 \right\} \\ &= E \left\{ [x - E(x)]^2 + 2[x - E(x)][y - E(y)] + [y - E(y)]^2 \right\} \\ &= E \left\{ [x - E(x)]^2 \right\} + E \left\{ [y - E(y)]^2 \right\} + 2E \left\{ [x - E(x)][y - E(y)] \right\}. \end{aligned}$$

• $\text{var}(x + y) = \text{var}(x) + \text{var}(y) + 2 \text{cov}(x, y)$. Q. E. D.

Theorem IV: $\text{var}(x - y) = \text{var}(x) + \text{var}(y) - 2 \text{cov}(x, y)$.

The proof follows the same method as for Theorem III.

The variance of the product or quotient of two variates depends on more complicated properties of the distribution functions. However, when the standard deviations of the distributions concerned are small compared to their expected values, it is possible to approximate the variance of a product as follows:

Theorem V: $\text{var}(xy) \approx \overline{E(y)}^2 \text{var}(x) + \overline{E(x)}^2 \text{var}(y) + 2E(x)E(y)\text{cov}(x, y)$.

$$\begin{aligned} \text{var}(xy) &= \text{var} \left\{ [E(x) + x - E(x)][E(y) + y - E(y)] \right\} \\ &= \text{var} \left\{ E(x)E(y) \left[1 + \frac{x - E(x)}{E(x)} \right] \left[1 + \frac{y - E(y)}{E(y)} \right] \right\} \\ &= E(x)^2 E(y)^2 \text{var} \left\{ 1 + \frac{x - E(x)}{E(x)} + \frac{y - E(y)}{E(y)} + \frac{[x - E(x)][y - E(y)]}{E(x)E(y)} \right\} \end{aligned}$$

or by equation III.42

$$\text{var}(xy) = E(x)^2 E(y)^2 \text{var} \left\{ \frac{x - E(x)}{E(x)} + \frac{y - E(y)}{E(y)} + \frac{[x - E(x)][y - E(y)]}{E(x)E(y)} \right\}.$$

Since the distribution is assumed to be narrow, we may neglect the last term to give

$$\begin{aligned} \text{var}(xy) &\approx E(x)^2 E(y)^2 \text{var} \left[\frac{x - E(x)}{E(x)} + \frac{y - E(y)}{E(y)} \right] \\ &\approx E(x)^2 E(y)^2 \text{var} \left[\frac{x}{E(x)} + \frac{y}{E(y)} \right]. \end{aligned}$$

Using Theorem III and equation III.43,

$$\begin{aligned} \text{var}(xy) &\approx E(x)^2 E(y)^2 \left[\frac{1}{E(x)^2} \text{var}(x) + \frac{1}{E(y)^2} \text{var}(y) + \frac{2}{E(x)E(y)} \text{cov}(x, y) \right] \\ &\approx E(y)^2 \text{var}(x) + E(x)^2 \text{var}(y) + 2E(x)E(y) \text{cov}(x, y). \quad \text{Q. E. D.} \end{aligned}$$

In the case where the two are independent, then we have

Theorem Va: If the two variates are independent in the statistical sense, then there is an exact relationship:

$$\text{var}(xy) = \overline{E(y)}^2 \text{var}(x) + \overline{E(x)}^2 \text{var}(y) + \text{var}(x) \cdot \text{var}(y)$$

$$\text{var}(xy) = E(x^2 y^2) - \overline{E(xy)}^2$$

Since x and y are independent, x^2 and y^2 are independent and consequently

$$E(x^2 y^2) = E(x^2) \cdot E(y^2)$$

$$E(xy) = E(x) \cdot E(y)$$

$$\begin{aligned} \therefore \text{var}(xy) &= E(x^2) \cdot E(y^2) - \overline{E(x)}^2 \overline{E(y)}^2 \\ &= [E(x)^2 + \text{var}(x)][E(y)^2 + \text{var}(y)] - E(x)^2 \cdot E(y)^2 \\ &= \overline{E(y)}^2 \text{var}(x) + \overline{E(x)}^2 \text{var}(y) + \text{var}(x) \cdot \text{var}(y). \end{aligned}$$

Q. E. D.

$$\text{Theorem VI: } \text{var} \left(\frac{x}{y} \right) \approx \frac{E(x)^2 \left[\frac{\text{var}(x)}{E(x)^2} + \frac{\text{var}(y)}{E(y)^2} - \frac{2 \text{cov}(x, y)}{E(x)E(y)} \right]}{E(y)^2}$$

$$\begin{aligned} \text{var} \left(\frac{x}{y} \right) &= \text{var} \left[\frac{E(x) + x - E(x)}{E(y) + y - E(y)} \right] \\ &= \text{var} \left[\frac{E(x)}{E(y)} \frac{1 + \frac{x - E(x)}{E(x)}}{1 + \frac{y - E(y)}{E(y)}} \right] \end{aligned}$$

$$\approx \left[\frac{E(x)}{E(y)} \right]^2 \text{var} \left\{ \left[1 + \frac{x - E(x)}{E(x)} \right] \left[1 - \frac{y - E(y)}{E(y)} \right] \right\}$$

since we have assumed that $\frac{y - E(y)}{E(y)}$ is small. The proof then follows as for Theorem V.

Finally, let us consider a to be a vector with constant components a_i and x to be a vector with the same number of variate components x_i ; then $a^t x$ is a linear combination of the components of x . We can define a covariance matrix of x , K_x , such that

$$(K_x)_{ij} = \text{cov}(x_i, x_j). \tag{III.44}$$

Theorem VII: $\text{Var}(a^t x) = a^t K_x a$ (in matrix notation).

$$\begin{aligned} a^t x &= \sum_i a_i x_i \\ \text{var} \left(\sum_i a_i x_i \right) &= E \left\{ \left[\sum_i a_i x_i - E \left(\sum_i a_i x_i \right) \right]^2 \right\} \\ &= E \left\{ \left[\sum_i a_i x_i - \sum_i a_i E(x_i) \right]^2 \right\} \\ &= E \left\{ \left[\sum_i a_i (x_i - E(x_i)) \right]^2 \right\} \\ &= E \left\{ \sum_i \sum_j a_i a_j [x_i - E(x_i)] [x_j - E(x_j)] \right\} \\ &= \sum_i \sum_j a_i a_j \text{cov}(x_i, x_j) \\ &= \sum_i \left\{ a_i \sum_j [\text{cov}(x_i, x_j) a_j] \right\} = a^t K_x a \text{ in matrix notation.} \end{aligned}$$

Q. E. D.

Theorem VII gives us the means of calculating the variance of any linear combination of variates with constant coefficients if the covariances of all pairs of components are known.

4. Averaging and Sampling Statistics

Let us consider first a large number m of independent quantities x_i which constitute a distribution. There is no basis for selection among these quantities other than their relation to each other and to the distribution defined by the measurements themselves. The expected value of x from this distribution is given by

$$E(x) = \frac{\sum_{i=1}^m x_i}{m} \quad (\text{III.45})$$

which is a simple arithmetic average, since the "probability of x having the value x_i " is constant for each observed value x_i . We shall symbolize $E(x)$ by X for convenience. Also

$$\text{var}(x) = \frac{\sum_{i=1}^m (x_i - X)^2}{m} = \sigma_x^2. \quad (\text{III.46})$$

Now let us select a subgroup or sample of n_j of the measurements chosen at random, so that $n_j < m$. The average of these n_j measurements will be, for the j th sample,

$$\bar{x}_j = \frac{\sum_{i=1}^{n_j} x_i}{n_j} = \sum_{i=1}^{n_j} \left(\frac{1}{n_j} x_i \right) \quad (\text{III.47})$$

(where the summation is taken over the n_j selected values, not necessarily the first n_j). By Theorem III, recalling that the measurements are all independent,*

$$\text{var}(\bar{x}_j) = \sum_{i=1}^{n_j} \text{var}\left(\frac{1}{n_j} x_i\right), \quad (\text{III.48})$$

and by III.41,

$$\text{var}(\bar{x}_j) = \sum_{i=1}^{n_j} \frac{1}{n_j^2} \text{var}(x) = \frac{1}{n_j^2} \sum_{i=1}^{n_j} \sigma_x^2, \quad (\text{III.49})$$

$$\text{var}(\bar{x}_j) = \frac{1}{n_j^2} (n_j \sigma_x^2), \quad (\text{III.50})$$

$$\text{var}(\bar{x}_j) = \frac{1}{n_j} \sigma_x^2, \quad (\text{III.51})$$

or, in other words, the variance of the average of a sample of n_j measurements is $1/n_j$ the variance of an individual measurement. This result can be used to describe the precision with which the true mean of the total distribution may be estimated from the mean of a random sample drawn from the distribution, provided the variance of the distribution of the sample can be obtained.

* This is strictly true only for an infinitely large population, or if each measurement is replaced in the population after selection so that it may be selected again.

This variance may be estimated from the distribution of the quantities within the sample, as follows: for a given quantity x_i in the j th sample

$$(x_i - X) = (x_i - \bar{x}_j) + (\bar{x}_j - X), \quad (\text{III.52})$$

$$(x_i - X)^2 = (x_i - \bar{x}_j)^2 + (\bar{x}_j - X)^2 + 2(x_i - \bar{x}_j)(\bar{x}_j - X). \quad (\text{III.53})$$

Summing over the total sample,

$$\sum_i^{n_j} (x_i - X)^2 = \sum_i^{n_j} (x_i - \bar{x}_j)^2 + n_j (\bar{x}_j - X)^2 + 2(\bar{x}_j - X) \sum_i^{n_j} (x_i - \bar{x}_j), \quad (\text{III.54})$$

but

$$\sum_i^{n_j} (x_i - \bar{x}_j) = 0. \quad (\text{III.55})$$

Therefore

$$\sum_i^{n_j} (x_i - X)^2 = \sum_i^{n_j} (x_i - \bar{x}_j)^2 + n_j (\bar{x}_j - X)^2, \quad (\text{III.56})$$

$$E \left[\sum_i^{n_j} (x_i - X)^2 \right] = E \left[\sum_i^{n_j} (x_i - \bar{x}_j)^2 \right] + n_j E (\bar{x}_j - X)^2, \quad (\text{III.57})$$

but

$$E \left[\sum_i^{n_j} (x_i - X)^2 \right] = \sum_i^{n_j} E \left[(x_i - X)^2 \right] = n_j \sigma_x^2 \quad (\text{III.58})$$

and

$$E \left[(\bar{x}_j - X)^2 \right] = \text{var}(\bar{x}_j) = \frac{1}{n_j} \sigma_x^2 \quad (\text{III.59})$$

so that, inserting III.58 and III.59 in III.57, and rearranging,

$$E \left[\sum_i^{n_j} (x_i - \bar{x}_j)^2 \right] = n_j \sigma_x^2 - n_j \left(\frac{1}{n_j} \sigma_x^2 \right) = (n_j - 1) \sigma_x^2. \quad (\text{III.60})$$

Now if we let

$$s_j^2 = \frac{\sum_i^n (x_i - \bar{x}_j)^2}{n_j - 1}, \quad (\text{III.61})$$

then

$$E(s_j^2) = \sigma_x^2 \quad (\text{III.62})$$

and s_j^2 is an unbiased estimate of σ_x^2 . Note that when the variance of a total distribution is estimated from a sample of n_j quantities, the denominator is $n_j - 1$; when the variance of the distribution formed by the sample itself is required, the denominator is n_j .

To recapitulate:

$$\text{var}(x) \approx \frac{\sum_i^{n_j} (x_i - \bar{x}_j)^2}{n_j - 1} \quad (\text{III.63})$$

is an estimated measure of the precision of an individual measurement;

$$\text{var}(\bar{x}) \approx \frac{\sum_i^{n_j} (x_i - \bar{x}_j)^2}{n_j(n_j - 1)} \quad (\text{III.64})$$

is an estimated measure of the precision of the average of n_j measurements.

One can obtain the same result in another manner. Let us assume that we have a set of n independent observations x_i belonging to a population with variance σ^2 .

Let

$$u_{ij} = x_i - x_j \quad (\text{III.65})$$

then

$$\text{var}(u_{ij}) = \text{var}(x_i) + \text{var}(x_j) - 2 \text{cov}(x_i, x_j) = E(u_{ij}^2) - \overline{E(u_{ij})}^2, \quad (\text{III.66})$$

$$\text{var}(x_i) = \text{var}(x_j) = \sigma^2, \quad (\text{III.67})$$

$$\text{cov}(x_i, x_j) = \delta_{ij} \sigma^2. \quad (\text{III.68})$$

Now

$$E(u_{ij}) = E(x_i) - E(x_j) = 0 \quad (\text{III.69})$$

and therefore

$$\text{var}(u_{ij}) = 2(1 - \delta_{ij})\sigma^2 = E(u_{ij}^2) = E(x_i^2 + x_j^2 - 2x_i x_j) \quad (\text{III.70})$$

so that an estimate of σ^2 can be made from

$$\frac{x_i^2 + x_j^2}{2} - x_i x_j \approx (1 - \delta_{ij})\sigma^2. \quad (\text{III.71})$$

This estimate can be improved as far as possible from the available data by measuring over all possible pairs.

$$\sum_i^n \sum_j^n \left(\frac{x_i^2 + x_j^2}{2} - x_i x_j \right) = n \sum_i^n x_i^2 - \left(\sum_i^n x_i \right)^2 \approx \sigma^2 \sum_i^n \sum_j^n (1 - \delta_{ij}), \quad (\text{III.72})$$

$$n \sum_i^n x_i^2 - \left(\sum_i^n x_i \right)^2 \approx (n^2 - n)\sigma^2, \quad (\text{III.73})$$

so that

$$\sigma^2 \approx \frac{\sum_i^n x_i^2 - \frac{\left(\sum_i^n x_i \right)^2}{n}}{n-1} \approx \frac{\sum_i^n (x_i - \bar{x})^2}{n-1} \quad (\text{III.74})$$

is the best estimate of σ^2 available from the data.

The value of this approach lies in the development of the concept of "degrees of freedom" of the system. It can be seen that the term -1 in the denominator arises from the restriction that the members of all possible

pairs are not independent, i. e. , that $\text{cov}(x_i, x_j)$ is not always zero. It can be shown that in statistical estimation, the denominator in an expression like III.75 is given by the number of pieces of information (here n) minus the number of relationships between them. In this case we have one relationship; we are taking an average, so that \bar{x} is expressed in terms of the x_i . Later we shall see that the resolution of n -component systems, when m data points are available, will involve $m - n$ degrees of freedom.

Let us now consider a series of independent measurements of the same quantity, each with its own variance, as sometimes happens. We wish to make a best estimate of the true value of that quantity by averaging the results, but we feel intuitively that the more precise results should be given greater weight in the averaging process; we have also stated that averaging is properly a process which applies only within a given distribution, which is equivalent to saying that the components of the average should all have the same variance. This is not true in the case we are considering; however, since all the measurements purport to be measurements of the same quantity, it should be possible to put them all on the same basis.

We have already seen that measurements which are mean values of sets, or samples, of measurements have variances which are reduced from the variances of individual measurements. Let us consider that a series of k sample sets have been drawn from a parent universe with a variance σ^2 ; and let the number of measurements in each set vary from sample to sample. The means of such a series would then represent a series of data such as we are discussing.

In this case, for the j th set, from III.51

$$\text{var}(\bar{x}_j) = \frac{1}{n_j} \sigma^2. \quad (\text{III.75})$$

Consequently,

$$n_j = \frac{\sigma^2}{\text{var}(\bar{x}_j)} \quad (\text{III.76})$$

and in order to combine all the observations into one grand average, we take

$$\bar{X} = \frac{\sum_{j=1}^k n_j \bar{x}_j}{\sum_{j=1}^k n_j} = \frac{\sum_{j=1}^k \frac{\bar{x}_j}{\text{var}(\bar{x}_j)}}{\sum_{j=1}^k \frac{1}{\text{var}(\bar{x}_j)}} \quad (\text{III.77})$$

and

$$\text{var}(\bar{X}) = \frac{1}{\sum_{j=1}^k n_j} \sigma^2 = \frac{1}{\left[\sum_{j=1}^k \frac{1}{\text{var}(\bar{x}_j)} \right]} \quad (\text{III.78})$$

We therefore see that the proper "weighting factors" for averaging measurements of unlike variances derived from sampling differences are the reciprocals of the variances; it seems apparent that any set of measurements of unlike variances could be duplicated from the same universe by a set with unlike variances arising from sampling, so that the conclusion should be generally applicable.

It can easily be shown that a "weighted average" calculated using the reciprocals of the individual variances for weighting factors is the average having the least variance. If the individual quantities which are to be averaged are not uncorrelated, some complications are encountered, as follows:

Let the set of quantities to be averaged be X_j , and the weighting factor associated with each X_j be W_j .

Then

$$\bar{X} = \frac{\sum_j W_j X_j}{\sum_j W_j} \quad (\text{III.79})$$

In order to preserve complete independence among the W_j we shall not assume that they are normalized. We then have, from III.40,

$$\text{var}(\bar{X}) = \frac{1}{\left(\sum_j W_j \right)^2} \text{var} \left(\sum_j W_j X_j \right) \quad (\text{III.80})$$

which from Theorem VII gives us

$$\text{var}(\bar{X}) = \frac{1}{\left(\sum_j W_j\right)^2} \sum_k \sum_j W_k W_j \text{cov}(X_k, X_j). \quad (\text{III.81})$$

We require the set of W_j to be chosen to give the best value of \bar{X} , which we define to be the value with the least variance. Consequently the derivative of $\text{var}(\bar{X})$ with respect to each of the W_j , separately, must equal zero. Letting an arbitrary member of the set be W_ℓ we have

$$\frac{\partial}{\partial W_\ell} [\text{var}(\bar{X})] = 0 = \frac{1}{\left(\sum_j W_j\right)^4} \left[\left(\sum_j W_j\right)^2 \cdot 2 \sum_j W_j \text{cov}(X_j, X_\ell) - 2 \left(\sum_j W_j\right) \cdot \sum_j \sum_k W_j W_k \text{cov}(X_k, X_j) \right]. \quad (\text{III.82})$$

This gives

$$\sum_j W_j \text{cov}(X_j, X_\ell) = \frac{\sum_j \sum_k W_j W_k \text{cov}(X_j, X_k)}{\sum_j W_j}. \quad (\text{III.83})$$

The right-hand member of this equation does not depend on ℓ , the index of the arbitrarily chosen W_j , and hence may be represented by a constant C whose value is to be determined. We can put equation III.83 in matrix-vector form to give

$$K_x W = Ci \quad (\text{III.84})$$

where C is as described above, W is the vector whose components are W_j , i is a vector (of the same dimension as W) whose components are all unity, and K_x is the covariance matrix (see Theorem VII) of the X_j . We then have, directly,

$$W = CK_x^{-1} i. \quad (\text{III.85})$$

It is obvious from III.79 that multiplication of the set W_j by any arbitrary constant will not change the value of \bar{X} , so that C may have any value; it may if desired be chosen to make the W_j normalized.

The variance of \bar{X} can be obtained by substituting this value of W into

III.81, which gives in matrix notation (realizing that $\sum_j W_j = i^t W$)

$$\text{var}(\bar{X}) = \frac{W^t K_x W}{C^2 (i^t W)^2} = \frac{C i^t (K_x^{-1})^t \cdot K_x \cdot C K_x^{-1} i}{C^2 (i^t K_x^{-1} i)^2} \quad (\text{III.86})$$

Canceling like terms, and letting $K_x K_x^{-1}$ equal unity, we have

$$\text{var}(\bar{X}) = \frac{i^t (K_x^{-1})^t i}{(i^t K_x^{-1} i)^2}, \quad (\text{III.87})$$

but since K_x is symmetrical, K_x^{-1} must be symmetrical, and hence

$$\text{var}(\bar{X}) = \frac{1}{i^t K_x^{-1} i} \quad (\text{III.88})$$

In matrix notation,

$$\bar{X} = \frac{W^t X}{i^t W} = \frac{i^t K_x^{-1} X}{i^t K_x^{-1} i} \quad (\text{III.89})$$

In the simple case where the X_j are in fact uncorrelated, K_x is a diagonal matrix whose diagonal elements are the variances of the X_j ; the diagonal elements of the inverse of K_x are therefore the reciprocals of the variances of the X_j . Inserting these values into III.88 and III.89, we arrive at III.77 and III.78.

5. The Least-Squares Solution

If we are given a set of data, knowing that the measurements of the set are subject to random error, and if we are given an equation of known mathematical form containing a number of unknown parameters whose values are to be determined by fitting the equation to the data, what is the procedure which will give the "best" estimates of the parameters from the admittedly

imperfect data? We obviously wish to minimize the average deviation between the observed points and the calculated points in some manner. The average of the algebraic values of the deviations will not do, since a large negative deviation would counterbalance a large positive one; we might consider minimizing the average of the absolute values of the deviations. A better approach is to minimize the average of the squares of the deviations of the observed from the calculated points. We shall show that this approach leads to expected values of the parameters which are the true values. In practice it is enough to minimize the sum of the squares of the deviations, properly weighted.

We must first consider the question of the relative worth of the experimental points. If we have no information about the estimated precision of each point, we have no choice but to treat them all alike; furthermore, if we have reason to believe that the points are all equally precise, we have every reason to treat them alike. However, if we have an estimate of the precision of each point, and the estimates are not all equal, we shall see that the points should be weighted.

We will see that the appropriate weighting factors in this case are inversely proportional to the variances of the points. Let us assume that we have a set of m measurements y_j of known variance; and also that each point is to be fitted by an expression which is a linear combination of n known functions or quantities a_{ij} with unknown coefficients x_i which are constant over the total set of measurements. We know that if the equations are consistent (as they must be to represent a real case) and if $m \geq n$, solutions for the x_i are possible; let us assume $m > n$. We then have a set of m equations

$$y_j = \sum_{i=1}^n a_{ij} x_i \quad (\text{III.90})$$

In general any well-selected set of parameters x_i will yield a calculated set of values of y_j which differ from the observed values by a set of small quantities,

$$E_j = \sum_{i=1}^n a_{ij} x_i - y_j \quad (\text{III.91})$$

We wish to minimize S , the weighted sum of the squares of the E_j , with respect to each parameter x_k . Let us for the moment assume that the appropriate weighting factor is proportional to the reciprocal of $\text{var}(y_j)$. Then

$$S = \sum_{j=1}^m \left[\frac{\left(\sum_{i=1}^n a_{ij} x_i - y_j \right)^2}{\text{var}(y_j)} \right], \quad (\text{III.92a})$$

$$\frac{\partial S}{\partial x_k} = \sum_{j=1}^m \frac{2a_{kj} \left(\sum_{i=1}^n a_{ij} x_i - y_j \right)}{\text{var}(y_j)} = 0, \quad (\text{III.92b})$$

$$\sum_{j=1}^m \frac{\sum_{i=1}^n \left(a_{kj} a_{ij} x_i \right) - a_{kj} y_j}{\text{var}(y_j)} = 0, \quad (\text{III.92c})$$

or, reversing the order of summation,

$$\sum_{i=1}^n \left\{ \left[\sum_{j=1}^m \frac{a_{kj} a_{ij}}{\text{var}(y_j)} \right] x_i \right\} = \sum_{j=1}^m \frac{a_{kj} y_j}{\text{var}(y_j)} \quad (\text{III.92d})$$

for each value of k from 1 to n . These equations are referred to as the normal equations of the system. Their solutions are the values for the x_i which minimize the quantity S .

It can be seen that if the equations of the original set III.90 are weighted by the reciprocal of the square root of the variance of y_j , the normal equations can be put in matrix notation very simply. Let the matrix A , the vector X , and the vector Y be defined by

$$A_{ij} = \frac{a_{ij}}{[\text{var}(y_j)]^{1/2}}, \quad (\text{III.93a})$$

$$X_i = x_i, \tag{III.93b}$$

$$Y_j = \frac{y_j}{[\text{var}(y_j)]^{1/2}}. \tag{III.93c}$$

Then the equation III.90 can be written

$$AX = Y. \tag{III.94}$$

Note that A is not a square matrix since $m > n$. If we multiply both sides of the equation from the left by the transpose of the matrix A, we get

$$A^tAX = A^tY, \tag{III.95}$$

which by the rules of matrix algebra can be shown to be precisely equivalent to equation III.92d, including the proper weighting factor. The solution in matrix notation is then

$$X = (A^tA)^{-1} \cdot A^tY. \tag{III.96}$$

For the remaining discussion we shall assume that the equations are properly weighted at the start.

Kenney and Keeping² present the following demonstration that if the equations are so weighted the expected values of the x_i calculated from the normal equations III.95 are the true values of the system.

Let us suppose that the y_j of equation III.90 are distributed about their true values η_j with a common variance σ^2 (which will be true if the equations are properly weighted) and that they are independent. Let us write the normal equations III.95 in the form

$$BX = g \tag{III.97}$$

where the matrix B is equal to A^tA and the vector g is equal to A^tY . Then

$$X = B^{-1}g, \tag{III.98a}$$

$$x_j = \sum_k (B^{-1})_{jk} g_k, \tag{III.98b}$$

$$\therefore E(x_j) = \sum_k (B^{-1})_{jk} E(g_k). \tag{III.98c}$$

²J. F. Kenney and E. S. Keeping, The Mathematics of Statistics, Part Two (D. Van Nostrand Co., Inc., Princeton, 1951), 2nd Ed., P. 309 ff.

We obtain the expected value of g_k from the expected values of the y_j ,

$$E(g_k) = \sum_i a_{ki} E(y_i), \quad (\text{III.98d})$$

and the expected values of the y_i from the true values ξ_ℓ of the x_ℓ ,

$$E(y_i) = \sum_\ell a_{\ell i} \xi_\ell. \quad (\text{III.98e})$$

Substituting back into equation III.98c,

$$E(x_j) = \sum_k \sum_i \sum_\ell (B^{-1})_{jk} \cdot a_{ki} \xi_\ell \cdot a_{\ell i}. \quad (\text{III.98f})$$

Rearranging the order of summation,

$$E(x_j) = \sum_\ell \left\{ \xi_\ell \sum_k \left[(B^{-1})_{jk} \sum_i a_{ki} a_{\ell i} \right] \right\}, \quad (\text{III.98g})$$

but

$$\sum_i a_{ki} a_{\ell i} = B_{k\ell}, \quad (\text{III.98h})$$

$$\therefore E(x_j) = \sum_\ell \left[\xi_\ell \sum_k (B^{-1})_{jk} \cdot B_{k\ell} \right]. \quad (\text{III.98i})$$

Now by the definition of an inverse matrix and the laws of matrix multiplication,

$$\sum_k (B^{-1})_{jk} B_{k\ell} = \delta_{j\ell} = \begin{cases} 1, & j=\ell, \\ 0, & j \neq \ell. \end{cases} \quad (\text{III.98j})$$

$$\therefore E(x_j) = \sum_\ell \xi_\ell \delta_{j\ell} = \xi_j. \quad \text{Q. E. D.} \quad (\text{III.98k})$$

That is to say, the expected values of the x_j resulting from solution of the properly weighted normal equations are the true values for the system.

We also wish information concerning the variances of the x_j so obtained, and of the variance of any calculated y using a suitable set of a_i . For both these problems we need the covariance matrix K_x of the vector x (see the definition, equation III.41).

We shall show, following Scheffé,³ that the inverse of the matrix of the normal equation, multiplied by the common variance of the y , is the covariance matrix of x .

Let us define a vector V whose components are jointly distributed random variates all having finite variance. Let us further define the expected value of a matrix or vector to be the matrix or vector formed from the expected values of its components. Then we can write

$$K_v = E \left\{ [V - E(V)] [V - E(V)]^t \right\}. \quad (\text{III.99})$$

Now if we have a vector W with variate components, and a matrix A with constant components, and if

$$W = AV, \quad (\text{III.100})$$

then

$$K_W = E \left\{ [W - E(W)] [W - E(W)]^t \right\}, \quad (\text{III.101a})$$

$$K_W = E \left\{ A[V - E(V)] [V - E(V)]^t A^t \right\}, \quad (\text{III.101b})$$

$$K_W = AE \left\{ [V - E(V)] [V - E(V)]^t \right\} A^t, \quad (\text{III.101c})$$

or

$$K_W = AK_v A^t. \quad (\text{III.102})$$

The normal equations are written,

$$A^t AX = A^t Y, \quad (\text{III.95})$$

and we assume that the covariance matrix of Y is

$$K_y = \sigma^2 I \quad (\text{III.103})$$

since the y_j are all independent and have a common variance σ^2 .

³H. Scheffé, The Analysis of Variance (John Wiley & Sons, Inc., New York 1959), pp. 8-12.

$$X = (A^t A)^{-1} A^t Y \quad \text{(III.104a)}$$

and from III.102

$$K_x = (A^t A)^{-1} A^t \cdot K_y \cdot \left[(A^t A)^{-1} A^t \right]^t \quad \text{(III.104b)}$$

$$\therefore K_x = \sigma^2 (A^t A)^{-1} \cdot A^t \cdot I \cdot A \cdot \left[(A^t A)^{-1} \right]^t \quad \text{(III.104c)}$$

But since $A^t A$ is symmetrical, $(A^t A)^{-1}$ is symmetrical and

$$\left[(A^t A)^{-1} \right]^t = (A^t A)^{-1} \quad \text{(III.104d)}$$

$$\therefore K_x = \sigma^2 (A^t A)^{-1} \cdot A^t A \cdot (A^t A)^{-1} \quad \text{(III.104e)}$$

$$\therefore K_x = \sigma^2 (A^t A)^{-1} \quad \text{Q. E. D.} \quad \text{(III.104f)}$$

We now need an unbiased estimate of σ^2 . We define the residual p_j by

$$p_j = y_j - Y_j \quad \text{(III.105)}$$

where y_j is the (weighted) observed value of y and Y_j is the (weighted) value computed from the least-squares solution values of the parameters x . We shall show that, if we have n parameters and m points,

$$\sigma^2 = \frac{E\left(\sum_{j=1}^m p_j^2\right)}{m - n} \quad \text{(III.106)}$$

following the treatment of Kenney and Keeping.⁴ We define the error Δ_j by

$$\Delta_j = y_j - \eta_j \quad \text{(III.107)}$$

where η_j is the true value of y_j .

We have

$$E(y_j^2) = E\left[(\eta_j + \Delta_j)^2\right] = \eta_j^2 + \sigma^2 \quad \text{(III.108)}$$

since $E(\Delta_j)$ is zero. If the true values of the x_i are ξ_i , as before,

$$\eta_j = \sum_i \xi_i a_{ij} \quad \text{(III.109)}$$

⁴ Op. cit., pp. 311-312.

$$\therefore E(y_j^2) = \sum_{i=1}^n \sum_{k=1}^n \xi_i \xi_k a_{ij} a_{kj} + \sigma^2. \quad (\text{III.110})$$

Summing over all j,

$$\sum_{j=1}^m E(y_j^2) = E\left(\sum_{j=1}^m y_j^2\right) = m\sigma^2 + \sum_{i=1}^n \sum_{k=1}^n \xi_i \xi_k \left(\sum_{j=1}^m a_{ij} a_{kj}\right), \quad (\text{III.111})$$

$$\sum_{j=1}^m a_{ij} a_{kj} = B_{ik} \quad \text{where } B = A^t A. \quad (\text{III.112})$$

$$\therefore E\left(\sum_{j=1}^m y_j^2\right) = m\sigma^2 + \sum_{i=1}^n \sum_{k=1}^n \xi_i \xi_k B_{ik}. \quad (\text{III.113})$$

We write the normal equations as

$$BX = A^t AX = A^t y = G, \quad (\text{III.114})$$

$$G_i = \sum_{k=1}^n B_{ik} X_k, \quad (\text{III.115})$$

$$\sum_{i=1}^n X_i G_i = \sum_{i=1}^n \sum_{k=1}^n B_{ik} X_i X_k, \quad (\text{III.116})$$

$$E\left(\sum_{i=1}^n X_i G_i\right) = \sum_{i=1}^n \sum_{k=1}^n B_{ik} E(X_i X_k) = \sum_{i=1}^n \sum_{k=1}^n B_{ik} \left[\xi_i \xi_k + \text{cov}(X_i X_k)\right], \quad (\text{III.117})$$

$$E\left(\sum_{i=1}^n X_i G_i\right) = \sum_{i=1}^n \sum_{k=1}^n B_{ik} \left[\xi_i \xi_k + \sigma^2 (B^{-1})_{ik}\right], \quad (\text{III.118})$$

$$E\left(\sum_{i=1}^n X_i G_i\right) = \sum_{i=1}^n \sum_{k=1}^n B_{ik} \xi_i \xi_k + n\sigma^2, \quad (\text{III.119})$$

$$\text{since } \sum_{k=1}^n B_{ik} (B^{-1})_{ik} = \sum_{k=1}^n \delta_{ik} = 1.$$

$$\sum_{i=1}^n \sum_{k=1}^n \xi_i \xi_k B_{ik} = E \left(\sum_{i=1}^n X_i G_i \right) - n\sigma^2. \quad (\text{III.120})$$

Now substituting in III.85 and rearranging,

$$E \left(\sum_{j=1}^m y_j^2 \right) - E \left(\sum_{j=1}^m X_j G_j \right) = (m - n)\sigma^2. \quad (\text{III.121})$$

We have also

$$p = y - Y = y - AX, \quad (\text{III.122})$$

$$A^t p = A^t y - A^t AX = G - A^t AX = 0. \quad (\text{III.123})$$

$$\sum_{j=1}^m p_j^2 = p^t p = (y^t - Y^t)p = y^t p - Y^t p, \quad (\text{III.124})$$

$$Y^t p = X^t A^t p = X^t \cdot 0 = 0. \quad (\text{III.125})$$

$$\therefore \sum_{j=1}^m p_j^2 = y^t p = y^t (y - AX) = y^t y - G^t X. \quad (\text{III.126})$$

$$\therefore \sum_{j=1}^m p_j^2 = \sum_{j=1}^m y_j^2 - \sum_{j=1}^m X_j G_j. \quad (\text{III.127})$$

$$E \left(\sum_{j=1}^m p_j^2 \right) = E \left(\sum_{j=1}^m y_j^2 \right) - E \left(\sum_{j=1}^m X_j G_j \right). \quad (\text{III.128})$$

Substituting in equation III.128, and rearranging terms,

$$\sigma^2 = \frac{E \left(\sum_j p_j^2 \right)}{m - n} \quad (\text{III.129})$$

so that

$$S^2 = \frac{\sum_j p_j^2}{m - n} \quad (\text{III.130})$$

is an unbiased estimate of σ^2 . It should be borne in mind that S^2 is an estimate which corrects for variances which were not included in the weighting; consequently a result for S^2 of less than unity means either that, fortuitously, the data fit the calculated curve better than one may reasonably expect; or, alternatively, that the original error estimates were too large. In the first case one should not insert the calculated value of S^2 into equation III.104f, but should merely assume σ^2 to be unity; if however there is any reason to believe that the original error estimates were made by a non-objective method, the result obtained is essentially meaningless. One object of this monograph is to provide experimenters with a set of rules for making the best estimate of the errors actually involved in their measurements; if a subjective, conservative "guess" is used, not only is the resulting final estimate of the precision of the experiment meaningless, but useful information is sometimes lost. The common procedure of guessing the precision of experimental data is indefensible.

If one now wishes to know the precision of each of the parameters x_i in our least-squares solution, it can readily be obtained from the inverse matrix of the normal equations, since

$$x_i = x_i \pm \sqrt{S^2 [(A^t A)^{-1}]_{ii}}. \quad (\text{III.131})$$

A derivation exists (due to H. Henris and H. B. Levy of this laboratory) to show that the choice of weighting factors described by Eqs. III.93a, b and c is in fact that choice which results in minimum values for the components of the matrix K_x , and hence in the most precise estimate of the components of the vector X . The derivation follows:

We have the relationship

$$A_0 X_0 = Y_0 \quad (\text{III.132})$$

where A_0 , X_0 , and Y_0 represent matrices of "true" values. Suppose we have m observed components of Y and there are n unknown components of X , with $m > n$, and that the components of Y are subject to error.

We wish to estimate θ , which is a linear function of the x_i 's

$$\theta = c^t X \quad (\text{III.133})$$

where c^t is a given row vector. We wish to ask:

How can we estimate x so that our resulting estimate of θ will have minimum variance?

$$\text{var } \theta = c^t K_x c \quad (\text{see Theorem VII, page (39)}) \quad (\text{III.134})$$

where K is the covariance matrix for the estimate of x .

Let A and Y represent matrices of observed values

A is (m × n) Y is (m × 1)

$$A = A_0 + dA \quad E(dA) = 0 \quad (\text{III.135})$$

$$Y = Y_0 + dY \quad E(dy) = 0$$

Let us define our estimate of X as

$$X = PY \quad \text{where} \quad PA = I \quad \text{defines } P \quad (\text{III.136})$$

We need

$$K_X = E \left\{ (X - X_0)(X - X_0)^t \right\} \quad (\text{III.137})$$

$$A_0 X_0 = Y_0 \quad (\text{III.138})$$

$$[A - (dA)]X_0 = Y - (dY) \quad (\text{III.139})$$

$$AX_0 = Y - [(dY) - (dA)X] \quad (\text{III.140})$$

since

$$PA = I \quad (\text{III.141})$$

$$X_0 = PY - P[(dY) - (dA)X_0] \quad (\text{III.142})$$

$$X - X_0 = P[(dY) - (dA)X_0] \quad (\text{III.143})$$

$$K_X = E \left\{ (X - X_0)(X - X_0)^t \right\} = E \left\{ P[(dY) - (dA)X] [(dY) - (dA)X]^t P^t \right\} \quad (\text{III.144})$$

$$K_X = P E \left\{ [(dY) - (dA)X] [(dY) - (dA)X_0]^t \right\} P^t. \quad (\text{III.145})$$

Let

$$V \equiv E \left\{ [(dY) - (dA)X_0] [(dY) - (dA)X_0]^t \right\} \quad (\text{III.146})$$

V is an m × m symmetric, positive-semidefinite matrix, that is $a^t V a \geq 0$

$$K_X = PVP^t \quad (\text{III.147})$$

Since

$$P \quad A = I \quad (\text{III.148})$$

$m \times n \quad n \times n$

P must be a n × m matrix that can be written in the form $(TA)^{-1}T$ where T is an n × m matrix such that the product TA is a non-singular matrix.

Therefore T must be of rank n (A must also be of rank n). We can choose T so that var θ is a minimum.

Let t_{ij} represent the elements of the T matrix. Then for $\text{var } \theta$ to be a minimum

$$\frac{\partial(\text{var } \theta)}{\partial t_{ij}} = 0 \quad \text{for all } t_{ij}. \quad (\text{III.149})$$

Since

$$\text{var } \theta = c^t K_x c \quad (\text{III.150})$$

then

$$\frac{\partial(\text{var } \theta)}{\partial t_{ij}} \text{ will } = 0 \quad \text{only if } \frac{\partial K_x}{\partial t_{ij}} = 0, \quad (\text{III.151})$$

since K_x is square, symmetric, and positive - semidefinite. Since

$$K_x = PVP^t \quad (\text{III.152})$$

$$\frac{\partial K_x}{\partial t_{ij}} = \frac{\partial P}{\partial t_{ij}} VP^t + PV \frac{\partial P^t}{\partial t_{ij}} = 0. \quad (\text{III.153})$$

Since V is symmetric, positive-semidefinite, both

$$\frac{\partial P}{\partial t_{ij}} VP^t \quad \text{and} \quad PV \frac{\partial P^t}{\partial t_{ij}} \quad \text{must equal } 0. \quad (\text{III.154})$$

Consider

$$\frac{\partial P}{\partial t_{ij}} VP^t = 0 \quad (\text{III.155})$$

$$P = (TA)^{-1} T \quad (\text{III.156})$$

$$\frac{\partial P}{\partial t_{ij}} = (TA)^{-1} \left[\frac{\partial T}{\partial t_{ij}} - \frac{\partial T}{\partial t_{ij}} A(TA)^{-1} T \right] \quad (\text{III.157})$$

$$\frac{\partial P}{\partial t_{ij}} VP^t = (TA)^{-1} \left[\frac{\partial T}{\partial t_{ij}} - \frac{\partial T}{\partial t_{ij}} A(TA)^{-1} T \right] VT^t (A^t T^t)^{-1} = 0. \quad (\text{III.158})$$

Therefore

$$\frac{\partial T}{\partial t_{ij}} \left[VT^t - A(TA)^{-1} T VT^t \right] = 0. \quad (\text{III.159})$$

In order for the above to be = 0 for all t_{ij} it is necessary that $VT^t - A(TA)^{-1} T VT^t = 0$. Therefore,

$$A(TA)^{-1} T (VT^t) = (VT^t). \quad (\text{III.160})$$

Let
$$M = A(TA)^{-1} T \quad \text{an } m \times m \text{ matrix of rank } n. \quad (\text{III.161})$$

$$B = VT^t \quad \text{an } m \times n \text{ matrix.} \quad (\text{III.162})$$

Consider the matrix M:

M is an idempotent matrix of order m and rank n. It will have n characteristic roots $\lambda = 1$ and $m - n$ characteristic roots $\lambda = 0$. Therefore, there will be no more than n independent vectors z for which

$$Mz = z. \quad (\text{III.163})$$

We know

$$MB = B \quad (\text{III.164})$$

so that the n column vectors of B must constitute such a set of n independent eigenvectors associated with M ($\lambda = 1$).

But also

$$MA = A. \quad (\text{III.165})$$

Therefore the column vectors of A must also constitute such a set. Since there can only be n independent eigenvectors associated with $\lambda = 1$, the column vectors of B must be linear combinations of the column vectors of A and

$$B = AQ \quad \text{where } Q \text{ is an elementary matrix.} \quad (\text{III.166})$$

$$\begin{aligned} B &= VT^t = AQ \\ T^t &= V^{-1}AQ \\ T &= Q^t A^t V^{-1} \end{aligned} \quad (\text{III.167})$$

$$P = (TA)^{-1} T \quad (\text{III.168})$$

$$P = (Q^t A^t V^{-1} A)^{-1} Q^t A^t V^{-1} \quad (\text{III.169})$$

$$P = (A^t V^{-1} A)^{-1} (Q^t)^{-1} Q^t A^t V^{-1} \quad (\text{III.170})$$

$$P = (A^1 V^{-1} A)^{-1} A^1 V^{-1} \quad (\text{III.171})$$

This is identical to a weighted least squares method with the weighting matrix proportional to V^{-1} .

Note that the choice of Q has no effect on P so that the estimate of x and its variance is not dependent on Q.

The quantity χ^2 may be defined in terms of the unweighted residuals r_j by

$$\chi^2 = \sum_{j=1}^m \frac{r_j^2}{\text{var}(y_j)} \quad (\text{III.172})$$

If our weighting factors for the least-squares solution are chosen to correspond with the true values of the variances of the y_j , then obviously the value of σ^2 is one, and from III.130 we get

$$E(\chi^2) = m - n. \quad (\text{III.173})$$

It can be shown that

$$\text{var}(\chi^2) = 2(m - n), \quad (\text{III.174})$$

and tables exist^{5,6} showing the probability of occurrence through pure random error of any deviation $\chi^2 - E(\chi^2)$, measured in terms of $[\text{var}(\chi^2)]^{1/2}$. The calculated value of χ^2 may then be used as a test of the goodness of fit of the assumed equations to the data. If the calculated value of χ^2 exceeds the expected value by an amount too great to attribute to random error, one suspects at least one of three difficulties: first, an error in the calculation; second, an improper estimate of the variances of the y_j ; or third, an improper form for the function being fitted to the data. If the first source of trouble is eliminated through careful checking, a choice can frequently be made between the other two on the following basis: if the known functions a_{ij} of the equation which is being fitted to the data are themselves smoothly varying functions of a precisely known parameter such as the time of the observation, then the residuals should be tabulated in the same order as the corresponding values of that parameter. If any systematic trend then exists in the residuals, the form of the equation is probably incorrect; if the residuals are scattered at random, the estimates of variance used for the weighting factors are too small.

In many cases, and specifically in the analysis of radioactive decay data, the number of terms in the equation is sometimes uncertain. It is desirable to determine whether the presence of a particular term or set of terms causes a significant improvement in the fit of the equation to the data. A test for this

⁵ R. A. Fisher and F. Yates, Statistical Tables (Hafner, New York, 1948).

⁶ E. S. Pearson and H. O. Hartley, Biometrika Tables for Statisticians (University Press, Cambridge, 1957).

exists if the measurements are independent. If we let

$$V = \sum_{j=1}^m \frac{r_j^2}{\text{var}(y_j)} \quad (\text{III. 175})$$

and we let V_{N_1} be the value of V when an equation of N_1 terms is fitted to m data points, and V_{N_2} be the value of V when an equation of N_2 terms is fitted, and $N_1 > N_2$;

$$F = \frac{V_{N_2} - V_{N_1}}{V_{N_1}} \cdot \frac{n - N_1}{N_1 - N_2} \quad (\text{III. 176})$$

A value of F significantly greater than one is statistical evidence that the $N_1 - N_2$ additional terms were more than fortuitously helpful. Tables giving the significance of values of F in terms of probability of occurrence due to pure chance are given in many collections.^{5,6} If the exact form of the additional terms is not known, unfortunately, this approach cannot be used.

Von Holdt⁷ has published a computational method for use with digital calculators having a limited number of significant figures. Such calculators may lose information during the course of the calculation, since the differences involved in computing determinants, or inverse matrices, are sometimes quite small compared to the numbers involved; if the difference between two numbers n_1 and n_2 on a decimal calculator is less than $10^{-p} n_1$, where p is the number of figures the calculator will hold, it will be "rounded" to zero and all information will be lost. The number of significant figures retained in the calculation $n_1 - n_2$ is of the order of $p - \log_{10} \frac{N_1}{N_1 - N_2}$. Von Holdt has shown that his computational method, although somewhat more laborious than more orthodox techniques, gives only half the figure loss; if one uses an eight-digit calculator, and the data are such that the ordinary method of solution gives a complete loss of information due to rounding errors, von Holdt's method will give results in which the rounding error is only one part in ten thousand. The inverse matrix of the normal equation is also obtained with less rounding error than that encountered when using more usual inversion techniques.

The technique is as follows: we write the weighted equations for the data in matrix form as

$$AX = Y \quad (\text{III. 177})$$

⁷ R. E. von Holdt, Proc. Western Computer Conf. (1955), p. 162.

where

$$A_{ij} = \frac{a_{ij}}{[\text{var}(y_j)]^{1/2}}, \quad (\text{III.178})$$

$$X_i = x_i, \quad (\text{III.179})$$

$$Y_j = \frac{y_j}{[\text{var}(y_j)]^{1/2}}. \quad (\text{III.180})$$

We then transform the matrix A as follows: to each column after the first add a suitable multiple k_{1i} of the first column so that the transformed columns are each orthogonal to the first column. This multiple is determined from the orthogonality requirement, i. e. ,

$$\sum_{j=1}^m a_{j1}(a_{ji} + a_{j1}k_{1i}) = 0, \quad (\text{III.181})$$

consequently

$$k_{1i} = -\frac{\sum_{j=1}^m a_{j1}a_{ji}}{\sum_{j=1}^m a_{j1}^2}. \quad (\text{III.182})$$

The matrix is transformed again in a similar fashion by adding multiples k_{2i} of the modified second column so that each column subsequent to the new second column is orthogonal to it; it is easy to show that orthogonality to the first column is not destroyed by this procedure. This transformation is continued until each column in turn has been used as a base, and all the columns of the matrix are mutually orthogonal.

It can be seen from the rules of matrix multiplication that the orthogonalization process is equivalent to multiplying the matrix A from the right by an upper triangular matrix U with unit diagonal elements to give a new matrix H so that

$$H = AU. \quad \therefore H^t = U^t A^t. \quad (\text{III.183})$$

The matrix U can be very simply obtained by treating a unit matrix of n columns by the same transformations as A, using the same multiplying factors k_{ji} .

Now writing the normal equation in the usual way, and remembering that $UU^{-1} = I$,

$$A^t A X = A^t Y, \quad (\text{III.184 a})$$

$$U^t A^t A U U^{-1} X = U^t A^t Y, \quad (\text{III.184 b})$$

$$H^t H U^{-1} X = U^t A^t Y. \quad (\text{III.184 c})$$

Since the columns of H are all mutually orthogonal,

$$H^t H = D \quad (\text{III.184 d})$$

where D is a diagonal matrix.

$$\therefore X = U D^{-1} U^t A^t Y \quad (\text{III.184 e})$$

which is our required solution. The reduction in rounding error arises from the fact that inverting a diagonal matrix involves nothing more than replacing each diagonal element by its reciprocal; there are no subtractions.

Furthermore, since

$$X = (A^t A)^{-1} A^t Y, \quad (\text{III.184 f})$$

we can see from III.184f and III.184e that

$$(A^t A)^{-1} = U D^{-1} U^t. \quad (\text{III.184 g})$$

Von Holdt⁷ also shows that if the matrix A is augmented, or expanded, by including an extra column equal to the vector Y, and this vector also is made orthogonal to each column in turn, it is transformed into a vector R such that

$$R_j = - \frac{r_j}{[\text{var}(y_j)]^{1/2}} \quad (\text{III.184 h})$$

or in other words, the components of R are the weighted residuals of the system, obtained even before the parameters are available to compute the points. It is then a simple matter to compute

$$\chi^2 = R^t R. \quad (\text{III.184 i})$$

His argument is geometrical: the matrix form of the original equation can be taken to be an attempt to express a vector Y by the linear combination of a set of vectors which are the columns of the matrix A. This set of vectors defines a subspace S of m-dimensional Euclidean space spanned by these vectors. Since the computed Y is a linear combination of these vectors it must lie in S; the vector R computed by orthogonalizing the vector of the observed Y is orthogonal to the m orthogonal vectors defining S and is therefore the component of Y (observed) orthogonal to S. Consequently,

$$R = Y (\text{observed}) - Y (\text{computed}) \quad (\text{III.185})$$

by vector subtraction in a space of m + 1 dimensions.

IV. STATISTICS OF COUNTING DATA

1. Nature of the Problems

In order to apply proper statistical calculations to counting data it is necessary to have information on the variances of such data. A number of authors have described the simpler statistical problems associated with the counting of radioactive samples, in general following the treatment of Rainwater and Wu.⁸ These authors discussed in considerable detail the following problems: a) If the expected value for a count is u , what is the probability $P_n(u)$ that exactly n counts will be observed? b) If exactly n counts are observed, what is the probability that the "true," or expected, value lies between u and $u + du$?

Our treatment of the above problems will in general follow that of Rainwater and Wu. We shall, however, endeavor to extend and refine their treatment somewhat by eliminating some approximations. We shall also consider briefly the statistical problems posed by the existence of counter background; and also the problem of the precision of the result of a measurement of a parameter of a nuclear reaction by counting a radioactive product.

2. Distribution Functions of Counting Results

Let us assume that in a sample of radioactive material there are N atoms present at the start of the counting period, and that each of them has the same probability p of disintegrating and producing a count during the counting interval. Probability p is the product of the probabilities: 1) that a given atom will disintegrate during the counting period; and 2) that an atom which disintegrates in the sample will produce a count in the detector. If there is only one species present, it is apparent from the integrated form of the fundamental law of radioactive decay (II.2) that if the number of the atoms in the sample at the start of the measurement is N , and the duration of the measurement is t , the expected number of atoms still present, i. e., the number which have not decayed, is $Ne^{-\lambda t}$; therefore the fraction of atoms expected to decay during the measurement is $(1 - e^{-\lambda t})$. This is the same as the probability that one atom will decay during the time t . The probability that an atom which disintegrates in the sample will produce a count in the detector is the counting efficiency of the counter, which of course depends on

⁸L. J. Rainwater and C. S. Wu, Nucleonics, October 1947, p. 60.

the decay scheme of the species in question, the geometry of the counting arrangement, and the radiation scattering and absorbing properties of the environment. Either or both of these two factors may be very small so that usually p is much less than 1. We now wish to calculate the probability P_n that precisely n counts will be recorded during the counting interval.

The probability that n consecutive atoms will produce counts is p^n , the probability that the remaining $N - n$ atoms will not produce counts is $(1 - p)^{N-n}$; and the probability of recording exactly n events is the sum of the products of these two probabilities for all possible groups of n atoms chosen from the sample N . The number of ways of choosing n atoms from a sample of N atoms is

$$C_N(n) = \frac{N!}{n!(N-n)!} \quad (\text{IV.1})$$

and therefore

$$P_n = p^n(1-p)^{N-n} \frac{N!}{n!(N-n)!} \quad (\text{IV.2})$$

This is known as the Bernoulli distribution.

Consider the binomial expansion

$$[p + (1 - p)]^N \equiv 1. \quad (\text{IV.3})$$

It can easily be seen that the terms of this expansion correspond to the values of P_n for all n from 0 to N , so that

$$\sum_{n=0}^N P_n = 1. \quad (\text{IV.4})$$

The expectation value of n is $E(n)$,

$$E(n) = \sum_{n=0}^N nP_n = \sum_{n=1}^N nP_n, \quad (\text{IV.5a})$$

$$E(n) = \sum_{n=1}^N n \frac{p^n(1-p)^{N-n}N!}{n!(N-n)!} = pN \sum_{n=1}^N \frac{p^{(n-1)}(1-p)^{N-n}}{(n-1)!(N-n)!} (N-1)! \quad (\text{IV.5b})$$

If we let $M = N - 1$ and $m = n - 1$, then $M - m = N - n$ and

$$E(n) = pN \sum_{m=0}^M \frac{p^m (1-p)^{M-m} M!}{m! (M-m)!} = pN \cdot 1, \quad (\text{IV.5c})$$

$$E(n) = pN, \quad (\text{IV.5d})$$

as one might anticipate. We can substitute in the Bernoulli distribution IV.2 to calculate the probability of observing n events, given p and the expected value $E(n)$.

In order to calculate the variance of n from its distribution function, we calculate

$$E(n^2) = \sum_{n=0}^N n^2 P_n = \sum_{n=1}^N n^2 P_n, \quad (\text{IV.6a})$$

and making the same substitutions,

$$E(n^2) = Np \sum_{m=0}^M (m+1) \frac{p^m (1-p)^{M-m} M!}{m! (M-m)!}, \quad (\text{IV.6b})$$

$$E(n^2) = Np \left[\sum_{m=0}^M m \cdot \frac{p^m (1-p)^{M-m} M!}{m! (M-m)!} + \sum_{m=0}^M \frac{p^m (1-p)^{M-m} M!}{m! (M-m)!} \right], \quad (\text{IV.6c})$$

which by comparison with IV.4, IV.5b, and IV.5d gives

$$E(n^2) = Np(Mp + 1) = Np \left[Np + (1-p) \right]. \quad (\text{IV.6d})$$

Recalling that

$$\text{var}(n) = E(n^2) - E(n)^2, \quad (\text{IV.7})$$

$$\text{var}(n) = Np(1-p) = (1-p) E(n), \quad (\text{IV.8})$$

so that the standard deviation σ_n of n is given by

$$\sigma_n = \sqrt{(1-p) E(n)}. \quad (\text{IV.9})$$

Since p is usually very small, it can usually be neglected.

The Bernoulli distribution function given by equation IV.2 can be considerably simplified in form in those cases where N is very large, p is very small, and n is very much less than N ; fortunately this is often true. We make the approximations

$$\frac{N!}{(N-n)!} = N(N-1)(N-2)\dots(N-n+1) \approx N^n, \quad (\text{IV.10})$$

$$(1-p)^{N-n} \approx e^{-p(N-n)} \approx e^{-pE(n)}. \quad (\text{IV.11})$$

Inserting these and equation IV.5d into equation IV.2 we get what is known as the "Poisson distribution,"

$$P_n = \frac{\overline{E(n)}^n e^{-E(n)}}{n!}. \quad (\text{IV.12})$$

As might be expected from the approximations involved in its derivation, this distribution gives

$$\text{var}(n) = E(n), \quad (\text{IV.13a})$$

$$\therefore \sigma_n = \sqrt{E(n)}. \quad (\text{IV.13b})$$

For cases where the Poisson distribution is a valid approximation to the Bernoulli distribution, and n is quite large, the Poisson distribution may be well approximated by the Gaussian, or normal, distribution. This approximation is obtained by replacing $n!$ by Stirling's approximation, and by transforming the variable n to $x = [n - E(n)]$. We then have

$$P_x = \frac{[E(n)]^{E(n)+x} \cdot e^{-E(n)}}{[E(n)+x]!} = \frac{[E(n)]^{E(n)} \cdot [E(n)]^x e^{-E(n)}}{[E(n)]! [E(n)+1][E(n)+2]\dots[E(n)+x]}, \quad (\text{IV.14a})$$

$$P_x = \frac{[E(n)]^{E(n)} e^{-E(n)}}{[E(n)]!} \cdot \frac{[E(n)]^x}{[E(n)+1][E(n)+2]\dots[E(n)+x]}, \quad (\text{IV.14b})$$

$$[E(n)]! \approx \sqrt{2\pi E(n)} \cdot \left[\frac{E(n)}{e} \right]^{E(n)}, \quad (\text{IV.14c})$$

$$\therefore P_x = \frac{1}{\sqrt{2\pi E(n)}} \cdot \frac{1}{\left[1 + \frac{1}{E(n)}\right] \left[1 + \frac{2}{E(n)}\right] \dots \left[1 + \frac{x}{E(n)}\right]} \quad (\text{IV.14d})$$

Since n is large, $n - E(n)$ will be small compared to n , so that we can approximate

$$\left[1 + \frac{z}{E(n)}\right] \approx e^{z/E(n)} \quad \text{when} \quad |z| \leq |x|. \quad (\text{IV.14e})$$

Consequently,

$$\left[1 + \frac{1}{E(n)}\right] \left[1 + \frac{2}{E(n)}\right] \dots \left[1 + \frac{x}{E(n)}\right] \approx \exp \left[\sum_{m=1}^x m/E(n) \right] = e^{(x^2+x)/2E(n)}. \quad (\text{IV.14f})$$

Neglecting x compared to x^2 , we have finally

$$P_x = \frac{1}{\sqrt{2\pi E(n)}} e^{-x^2/2E(n)} \quad (\text{IV.14g})$$

or

$$P_x = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-x^2/2\sigma_x^2} \quad (\text{IV.14h})$$

which is the normal distribution with $\text{var}(x) = E(n)$.

3. Distribution of Expected Values for a Given Observed Count - Small Sampling

We can attack the inverse problem very easily. We now wish to know the expected value of n , and its variance, if we have an actual observed value n .

This is most simply approached for the case in which the Poisson distribution is applicable. We rewrite equation IV.12, letting the symbol u stand for $E(n)$ for convenience,

$$P_n(u) = \frac{u^n e^{-u}}{n!}, \quad (\text{IV.15})$$

to show that $P_n(u)$ is a function both of u and of n .

We now state the problem: for an actual count n , what is the probability that the actual value x of u lies between u and $u + du$? The question may be rephrased: if we consider repeating the identical experiment a large number of times, what is the predicted value of the average result given that the result of a single experiment is n ; and how reliable is the prediction, that is, what is its variance? In cases where the number of counts is large, it can be seen that equation IV.15 is the solution of the inverse problem when written

$$P_n(x) dx = \frac{x^n e^{-x}}{n!} dx. \quad (\text{IV.16})$$

From this we get the expected value of x from

$$E(x) = \int_0^{\infty} x P_n(x) dx = \frac{1}{n!} \int_0^{\infty} x^{n+1} e^{-x} dx = \frac{(n+1)!}{n!}, \quad (\text{IV.17a})$$

$$\therefore E(x) = n + 1. \quad (\text{IV.17b})$$

The fact that the expected value is larger than the observed value is surprising at first, until we consider that even if the observed value is zero, the expected value need not be zero.

Calculation of the variance of x gives

$$\text{var}(x) = \int_0^{\infty} x^2 P_n(x) dx - \overline{E(x)}^2, \quad (\text{IV.17c})$$

$$\text{var}(x) = \frac{(n+2)!}{n!} - (n+1)^2 = (n+2)(n+1) - (n+1)^2, \quad (\text{IV.17d})$$

$$\text{var}(x) = n + 1 = E(x). \quad (\text{IV.17e})$$

4. Distribution of Expected Values for a Given Observed Count - Large Sampling

Rainwater and Wu⁸ did not consider the problem of predicting expected count and count variance for cases where an appreciable fraction of the atoms present might be expected to give rise to counts. For such a case to arise, two conditions must be met simultaneously. First, the counting equipment must have a high efficiency for detection, which includes both a high solid angle for radiation from the sample to enter the counter and a high probability

that entering radiation will cause a pulse in the counting circuit. Second, the counting period must be at least an appreciable fraction of a half-life of the species involved. Such conditions are met in experiments involving short-lived species formed in low yield, where in order to obtain data of maximum reliability it is necessary to use high-efficiency counters and count for extended periods. Experiments designed to detect minimal amounts of certain contaminants may also be conducted under these conditions, where very long counts are resorted to.

Under these conditions the assumptions which lead to the Poisson distribution are no longer valid. We must therefore work with the more accurate but less tractable Bernoulli distribution. Our problem remains the same as in the previous section: Given an observed number of events n , what is the expected average number of events during a number of identical experiments, and what is the reliability of that number?

We approach the problem via a conceptual experiment. We assume that the experiment is carried out a very large number of times with a variety of different numbers of atoms of material. Let the number of times the experiment is performed be w and the probability that the number N of atoms in a given experiment has the value z be $\pi_N(z)$. Let w be so large that the number of experiments conducted in which N has the exact value z be $w\pi_N(z)$ as precisely as we choose.

Then if N has the value z , the probability that the number of counts observed, x , has the precise value n is given by the Bernoulli distribution.

$$P_x(n, z) = \frac{p^n (1-p)^{z-n} z!}{n! (z-n)!} \quad (\text{IV.18})$$

where p is, as before, the probability that a single atom will decay and give rise to a count during the course of the experiment.

The number of times this result will be observed in w experiments, if w is sufficiently large, is

$$m_x(n, z) = w\pi_N(z)P_x(n, z). \quad (\text{IV.19})$$

We know that if n events were observed there were at least n atoms in the sample. Therefore the total number of experiments, for any value of

N equal to or greater than n, which will give rise to a value n for x is

$$M_x(n) = \sum_{z=n}^{\infty} w\pi_N(z)P_x(n, z). \quad (IV.20)$$

Consequently, the fraction of the experiments leading to the value n for x in which N had the value z is given by the ratio of equations IV.19 and IV.20. If w is sufficiently large, this constitutes a definition of probability, so that

$$P_N(z) = \frac{w\pi_N(z)P_x(n, z)}{\sum_{z=n}^{\infty} w\pi_N(z)P_x(n, z)} \quad (IV.21)$$

or

$$P_N(z) = \frac{p^n(1-p)^{z-n} z! \pi_N(z)}{n!(z-n)! \sum_{z=n}^{\infty} \frac{p^n(1-p)^{z-n} z!}{n!(z-n)!} \pi_N(z)} \quad (IV.22)$$

or, simplifying,

$$P_N(z) = \frac{z!(1-p)^z}{(z-n)!} \cdot \frac{\pi_N(z)}{\sum_{z=n}^{\infty} \frac{z!(1-p)^z}{(z-n)!} \cdot \pi_N(z)}. \quad (IV.23)$$

In order to proceed further it is necessary to make an assumption. We assume, lacking any other information, that any value of N on our series of w experiments is as likely as any other; or alternatively, we assume that if we conduct our experiments in such a way that all values of N are equally likely, the conclusions drawn from such a set of experiments may be applied to actual laboratory results.

If we make such an assumption, the $\pi_N(z)$ are all equal and will cancel. In that case we are left with an expression

$$P_N(z) = \frac{z!(1-p)^z}{(z-n)! \sum_{z=n}^{\infty} \frac{z!(1-p)^z}{(z-n)!}} \quad (IV.24)$$

which is obviously normalized.

It is necessary to evaluate the sum of the infinite series in the denominator. The value of this series depends only on the constants n and p . Let us define the symbol

$$Q_n = \sum_{z=n}^{\infty} \frac{z! (1-p)^z}{(z-n)!} \quad (IV.25)$$

If n is zero, then

$$Q_0 = \sum_{z=0}^{\infty} (1-p)^z = 1 + (1-p) + (1-p)^2 + (1-p)^3 + \dots \quad (IV.26)$$

Since $1-p$ is less than one, this is a convergent geometric series with a ratio $1-p$, and

$$Q_0 = \frac{1}{1-(1-p)} = \frac{1}{p} \quad (IV.27)$$

Let us now consider Q_{m+1} , and let $y = z - 1$ or $z = y + 1$.

$$Q_{m+1} = \sum_{z=m+1}^{\infty} \frac{z! (1-p)^z}{[z-(m+1)]!} = \sum_{y=m}^{\infty} \frac{(y+1)! (1-p)^{y+1}}{(y+1-m-1)!} \quad (IV.28a)$$

$$Q_{m+1} = (1-p) \sum_{y=m}^{\infty} (y+1) \cdot \frac{y! (1-p)^y}{(y-m)!} \quad (IV.28b)$$

$$Q_{m+1} = (1-p) \left[\sum_{y=m}^{\infty} y \cdot \frac{y! (1-p)^y}{(y-m)!} + \sum_{y=m}^{\infty} \frac{y! (1-p)^y}{(y-m)!} \right] \quad (IV.28c)$$

$$Q_{m+1} = (1-p) \left[\sum_{y=m}^{\infty} (y-m+m) \frac{y! (1-p)^y}{(y-m)!} + Q_m \right] \quad (IV.28d)$$

$$Q_{m+1} = (1-p) \left[\sum_{y=m}^{\infty} (y-m) \cdot \frac{y! (1-p)^y}{(y-m)!} + m \sum_{y=m}^{\infty} \frac{y! (1-p)^y}{(y-m)!} + Q_m \right] \quad (IV.28e)$$

$$Q_{m+1} = (1 - p) \left[\sum_{y=m+1}^{\infty} (y - m) \frac{y! (1 - p)^y}{(y - m)!} + (m + 1) Q_m \right]. \quad (\text{IV.28f})$$

Note the change in the lower limit on the summation sign. This is allowable, since the first term when $y = m$ is zero. Dividing $(y - m)$, which is not zero, into the numerator and denominator in the summation,

$$Q_{m+1} = (1 - p) \left\{ \sum_{y=m+1}^{\infty} \frac{y! (1 - p)^y}{[y - (m + 1)]!} + (m + 1) Q_m \right\} \quad (\text{IV.28g})$$

or

$$Q_{m+1} = (1 - p) \left[Q_{m+1} + (m + 1) Q_m \right] \quad (\text{IV.28h})$$

which gives, immediately,

$$Q_{m+1} = \left(\frac{1 - p}{p} \right) (m + 1) Q_m \quad (\text{IV.28i})$$

and since $Q_0 = 1/p$ from equation IV.27,

$$Q_y = \left(\frac{1 - p}{p} \right)^y \cdot \frac{y!}{p}. \quad (\text{IV.28j})$$

It should be noted that equation IV.28c can be transformed into a useful form:

$$\sum_{y=m}^{\infty} y \cdot \frac{y! (1 - p)^y}{(y - m)!} = \frac{Q_{m+1}}{1 - p} - Q_m = \frac{m + 1 - p}{p} Q_m. \quad (\text{IV.29})$$

We now wish to get the expected value and variance of the number of atoms in the sample, as well as the expected value and variance of the number of counts in the counter. The expected value of N is

$$E(N) = \sum_{z=n}^{\infty} z P_N(z) = \sum_{z=n}^{\infty} z \cdot \frac{z! (1 - p)^z}{(z - n)! Q_n}, \quad (\text{IV.30})$$

substituting Q_n for the equivalent summation in equation IV.24. We then have, substituting in equation IV.30 from equation IV.29,

$$E(N) = \frac{1}{Q_n} \sum_{z=u}^{\infty} z \cdot \frac{z!(1-p)^z}{(z-n)!} = \frac{1}{Q_n} \left(\frac{n+1-p}{p} \cdot Q_n \right), \quad (\text{IV.31})$$

$$E(N) = \frac{n+1-p}{p}. \quad (\text{IV.32})$$

The variance of N is obtained from the relationship

$$\text{var}(N) = E(N^2) - \overline{E(N)}^2. \quad (\text{IV.33a})$$

$$E(N^2) = \sum_{z=n}^{\infty} z^2 P_N(z) = \frac{1}{Q_n} \sum_{z=n}^{\infty} z^2 \frac{z!(1-p)^z}{(z-n)!}, \quad (\text{IV.33b})$$

$$E(N^2) = \frac{1}{Q_n} \sum_{z=n}^{\infty} \left[z^2 - (2n+1)z + n(n+1) + (2n+1)z - n(n+1) \right] \frac{z!(1-p)^z}{(z-n)!}, \quad (\text{IV.33c})$$

$$E(N^2) = \frac{1}{Q_n} \sum_{z=n}^{\infty} (z-n)[z-(n+1)] \frac{z!(1-p)^z}{(z-n)!} + \frac{(2n+1)}{Q_n} \sum_{z=n}^{\infty} z \frac{z!(1-p)^z}{(z-n)!} - n(n+1) \quad (\text{IV.33d})$$

$$E(N^2) = \frac{1}{Q_n} \sum_{z=n+2}^{\infty} (z-n)[z-(n+1)] \cdot \frac{z!(1-p)^z}{(z-n)!} + (2n+1) \left(\frac{n+1-p}{p} \right) - n(n+1), \quad (\text{IV.33e})$$

$$E(N^2) = \frac{1}{Q_n} \sum_{z=n+2}^{\infty} \frac{z!(1-p)^z}{[z-(n+2)]!} + \left[\frac{(2n+1)(n+1)}{p} - (n^2 + 3n + 1) \right], \quad (\text{IV.33f})$$

$$E(N^2) = \frac{Q_{n+2}}{Q_n} + \left[\frac{(2n+1)(n+1)}{p} - (n^2 + 3n + 1) \right] = \left(\frac{1-p}{p} \right)^2 (n+2)(n+1) + \left[\frac{(2n+1)(n+1)}{p} - (n^2 + 3n + 1) \right], \quad (\text{IV.33g})$$

$$E(N^2) = \frac{(n+2)(n+1)}{p^2} - \frac{2(n+2)(n+1) - (2n+1)(n+1)}{p} + (n+2)(n+1) - n^2 - 3n - 1, \quad (\text{IV.33h})$$

$$E(N^2) = \frac{n+1}{p^2} + \frac{(n+1)^2}{p^2} - \frac{3(n+1)}{p} + 1. \quad (\text{IV.33i})$$

Therefore

$$\text{var}(N) = \frac{n+1}{p^2} + \left(\frac{n+1}{p}\right)^2 - \frac{3(n+1)}{p} + 1 - \left(\frac{n+1}{p} - 1\right)^2, \quad (\text{IV.34a})$$

$$\text{var}(N) = \frac{n+1}{p^2} - \frac{n+1}{p}, \quad (\text{IV.34b})$$

$$\text{var}(N) = \frac{(n+1)(1-p)}{p^2}. \quad (\text{IV.34c})$$

In order to determine $E(n)$, we recall that for the Bernoulli distribution

$$E(n) = E(pN) = pE(N). \quad (\text{IV.35})$$

$$E(n) = u = n+1 - p \quad (\text{IV.36})$$

which may be compared with equation IV.17b, derived from the Poisson distribution. It will be seen that the approximations involved in converting the Bernoulli distribution to the Poisson distribution result in the loss of the term $(-p)$; it is easy to see that if p approaches 1 (that is, if any atoms present are nearly certain to be counted) then the expected value of the count approaches the value of the count itself; for example, it is no longer likely that the expected value will be much greater than zero when an actual observation of zero is recorded.

The variance of n is calculated from

$$\text{var}(n) = \text{var}(pN) = p^2 \text{var}(N). \quad (\text{IV.37})$$

$$\text{var}(n) = (1-p)(n+1) \quad (\text{IV.38})$$

which differs from the variance deduced from the Poisson distribution by the factor $(1-p)$. It is obvious that when p is one (that is, that every atom present is certain to count), that the result is not subject to variation; hence its variance must be zero, as it is; the number of atoms in the sample can be measured with certainty.

5. Correction for Counter Background

It can be shown (see Rainwater and Wu⁸) that in cases where the Poisson distribution can be used it is not necessary to assume that p is constant for every atom; in other words the presence of many different species in the sample does not affect our statistical conclusions. If we consider the counter background to arise from a "species" with a very long half-life, its expected value and variance will be given by IV.17b and IV.17e. A word of caution is in order: if the counting equipment is malfunctioning in any of a number of ways to give rise to a nonrandom background, the results no longer apply, in particular as regards variance. If the background is random, then a measurement of background is independent of the gross count taken on the sample, in the statistical sense; consequently the covariance of the two measurements is zero, and the variance of a net count after subtracting the result of a background measurement is the sum of the variances of the background measurement and the gross count.

When the Poisson distribution is not applicable to the sample, the situation becomes more complicated. Fortunately, the Poisson distribution is always applicable to the background. It is necessary to take a background measurement, to deduce from this result the expected background count (which is the observed count plus one) and to calculate the probability of each possible value b of the background $P_B(b)$ from zero to g , where g is the gross count recorded for the sample. Each such probability is then normalized by the sum of all the probabilities $\sum_{b=0}^g P_B(b)$; they must be normalized from the

Poisson distribution since the background count cannot take on all possible values; negative values of the net count are not possible. If the number of events in the gross count, g , is considerably larger than the background count the normalizing factor will be very nearly one, but for counts not greatly in excess of background it is very different from one.

Each renormalized probability that the background count had the value b is now the probability that in the gross count the count X due to the sample had the value $g - b$, which may be written

$$\frac{1}{R_g} \cdot P_B(b) = P_x(g - b) \quad (\text{IV.39})$$

where $1/R_g$ is the renormalizing factor,

$$\frac{1}{R_g} = \frac{1}{\sum_{b=0}^g P_B(b)}, \quad (\text{IV.40})$$

and $P_B(b)$ is calculated from the Poisson distribution IV.12, with $u = E(n)$ taken as the observed number of events during the measurement, plus one, and n taken to be each value of b in turn.

We then calculate $E(N)$ and $\text{var}(N)$, or $E(n)$ and $\text{var}(n)$, as desired, from equations IV.32 and IV.34c or IV.36 and IV.38, for sample count by multiplying each expected value by $P_x(g-b)$ and summing over all values. Let the observed background count (over the same period) be β ; then

$$u = \beta + 1 = E(B) \quad (\text{IV.41})$$

and

$$R_g(u) = \sum_{j=0}^g \frac{u^j e^{-u}}{j!} \longrightarrow 1 \quad \text{as } g/u \longrightarrow \infty. \quad (\text{IV.42})$$

Then

$$E(N) = \sum_{b=0}^g \left(\frac{n+1-p}{p} \right) \frac{u^{g-n} e^{-u}}{(g-n)!} \cdot \frac{1}{R_g(u)}, \quad (\text{IV.43a})$$

$$E(N) = \sum_{b=0}^g \left(\frac{g+1-p}{p} - \frac{b}{p} \right) \frac{u^b e^{-u}}{b!} \cdot \frac{1}{R_g(u)}, \quad (\text{IV.43b})$$

$$E(N) = \frac{g+1-p}{p} - \frac{1}{p} \sum_{b=0}^g b \frac{u^b e^{-u}}{b!} \cdot \frac{1}{R_g(u)}, \quad (\text{IV.43c})$$

$$E(N) = \frac{g+1-p}{p} - \frac{1}{p} \sum_{b=1}^g \frac{u^b e^{-u}}{(b-1)!} \cdot \frac{1}{R_g(u)}. \quad (\text{IV.43d})$$

Let $y = 1 - b$, then

$$E(N) = \frac{g+1-p}{p} - \frac{u}{p} \cdot \frac{R_{g-1}(u)}{R_g(u)} = \frac{g+1-p-u}{p} + \frac{u}{p} \cdot \frac{u^g e^{-u}}{g! R_g(u)}$$

(IV.43e)

When the gross count is sufficiently above the background, the last term becomes very small and may be neglected. In this case

$$E(N) \approx \frac{g+1-p-u}{p} = \frac{g-\beta-p}{p}$$

(IV.44)

If the observed background count is taken over a period t_b of different length than the duration of the gross count t_g ,

$$E(N) \approx \frac{g+1-p - \frac{t_g}{t_b}(\beta+1)}{p}$$

(IV.45)

since we expect the background count to be uniformly distributed in time.

The variance of N is calculated from equation IV.33.

$$E(N^2) = \sum_{b=0}^g \left[\frac{n+1}{p^2} + \frac{(n+1)^2}{p^2} - \frac{3(n+1)}{p} + 1 \right] \frac{u^b e^{-u}}{b! R_g(u)},$$

(IV.46a)

$$E(N^2) = \sum_{b=0}^g \left[\frac{g-b+1}{p^2} + \frac{(g-b+1)^2}{p^2} - \frac{3(g-b+1)}{p} + 1 \right] \frac{u^b e^{-u}}{b! R_g(u)},$$

(IV.46b)

$$\text{var}(N) = E(N^2) - \overline{E(N)}^2.$$

(IV.46c)

Going through the algebra as before, and remembering that

$$\sum_{b=0}^g b \frac{u^b e^{-u}}{b!} = \sum_{b=1}^g b \frac{u^b e^{-u}}{b!} = \sum_{b=1}^g \frac{u^b e^{-u}}{(b-1)!} = u \sum_{(b-1)=0}^{g-1} \frac{u^{(b-1)} e^{-u}}{(b-1)!} = u R_{g-1}(u),$$

(IV.47a)

$$\sum_{b=0}^g b^2 \frac{u^b e^{-u}}{b!} = \sum_{b=1}^g b \cdot \frac{u^b e^{-u}}{(b-1)!} = \sum_{(b-1)=0}^{g-1} \left[(b-1) + 1 \right] \frac{u^{b-1+1} e^{-u}}{(b-1)!}$$

$$= u R_{g-1}(u) + u^2 R_{g-2}(u),$$

(IV.47b)

we finally get, after a certain amount of algebra,

$$\text{var}(N) = \frac{(1-p)(g+1)}{p^2} + \frac{u}{p} \frac{R_{g-1}(u)}{R_g(u)} + \frac{u^2}{p^2} \left[\frac{R_{g-2}(u)}{R_g(u)} - \frac{R_{g-1}(u)}{R_g(u)} \right]^2. \quad (\text{IV.48})$$

If, as before, $g \gg u$,

$$\text{var}(N) \approx \frac{(1-p)(g+1) + up}{p^2}. \quad (\text{IV.49})$$

As before, when $g \gg u$,

$$E(n) = E(pN) = pE(N) \approx g - \beta - p, \quad (\text{IV.50a})$$

$$\text{var}(n) = \text{var}(pN) = p^2 \text{var}(N) \approx (1-p)(g+1) + up = (1-p)(g+1) + p(\beta+1). \quad (\text{IV.50b})$$

When g is not much greater than u , the $R_g(u)$, $R_{g-1}(u)$, and $R_{g-2}(u)$ must be evaluated and the full equations derived from equations IV.43e and IV.48 must be used. This procedure is extremely laborious. Fortunately, we are frequently concerned with the parameters of the reaction forming the active species, rather than with the actual numbers of atoms in the sample - as will be shown, the Poisson distribution is always applicable in such a case.

6. Errors of Extrapolation

It rarely happens that counting can be started immediately after the creation of the radioactive species. If some delay occurs between the time of interest to the experimenter, t_0 , and the time of the measurement, t , one can extrapolate back to t_0 by multiplying by the exponential factor $e^{\lambda(t-t_0)}$. It is important to realize, however, that no matter how precisely the counting rate at t is known, and no matter how precisely the half-life or decay constant is known, a statistical uncertainty is introduced by the extrapolation itself. This is readily seen by considering that the atoms still present at t are those which have not decayed in the interval t_0 to t . Since those which have decayed are subject to statistical fluctuations, the sum of those still present and those which have decayed is subject to such fluctuation.

The variance of the extrapolated number can be derived as follows: Let the number of atoms present at t_0 be N_0 , the number of atoms decaying in

the interval t_0 to t be n , the number of atoms remaining at time t be N ; the extrapolation factor is then N_0/N .

Now the probability p of one atom decaying in the interval t_0 to t is $1 - e^{-\lambda(t-t_0)}$ where λ is the decay constant of the species.

The probability that n atoms decay during the interval t_0 to t is obtained from the Bernoulli distribution, since p is not necessarily much smaller than 1. We have, therefore,

$$P_n = p^n (1 - p)^{N_0 - n} \cdot \frac{N_0!}{n_0! (N_0 - n)!} \quad (\text{IV.51})$$

as the probability that n atoms will decay during the interval of interest.

Substituting

$$\text{a) } n = N_0 - N, \quad (\text{IV.52a})$$

$$\text{b) } p = 1 - e^{-\lambda(t-t_0)}, \quad (\text{IV.52b})$$

we get

$$P_n = \left[1 - e^{-\lambda(t-t_0)} \right]^{N_0 - N} \cdot \left(e^{-\lambda t} \right)^N \cdot \frac{N_0!}{(N_0 - N)! \cdot N!} \quad (\text{IV.53})$$

Now the probability that n atoms decay is the same as the probability that N atoms remain, so that

$$P_N = P_n \quad (\text{IV.54})$$

The form of the distribution function for N is then seen to be a Bernoulli distribution with a probability $p' = e^{-\lambda(t-t_0)}$, and from equations IV.32 and IV.34c,

$$E(N_0) = \frac{N + 1 - p'}{p'} = (N + 1) e^{\lambda(t-t_0)} - 1, \quad (\text{IV.55a})$$

$$\text{var}(N_0) = \frac{(N + 1)(1 - p')}{(p')^2} = (N + 1) \left[1 - e^{-\lambda(t-t_0)} \right] \cdot e^{2\lambda(t-t_0)}. \quad (\text{IV.55b})$$

The extrapolation factor is given by

$$F_e = E(N_0)/N = \left[\frac{N + 1 - e^{-\lambda(t-t_0)}}{N} \right] e^{\lambda(t-t_0)} = \frac{1 - e^{-\lambda(t-t_0)}}{N} + e^{\lambda(t-t_0)} \quad (\text{IV.56})$$

which for N large compared to 1 is simply $e^{\lambda(t-t_0)}$, which is as expected; only for extremely small values of N is there an appreciable correction. The variance of the extrapolation factor is given by

$$\text{var}(F_e) = \frac{1}{N^2} \text{var}(N_0) = \frac{(N + 1) \left[1 - e^{-\lambda(t-t_0)} \right] e^{2\lambda(t-t_0)}}{N^2} \quad (\text{IV.57})$$

which for values of N large compared to 1 simplifies to

$$\text{var}(F_e) \approx \left[\frac{1 - e^{-\lambda(t-t_0)}}{N} \right] e^{2\lambda(t-t_0)}. \quad (\text{IV.58})$$

When extrapolating from a known number of atoms at time t , with known variance, back to a desired time t_0 , we have from Theorem V, Chapter III,

$$N_0 = NF_e, \quad (\text{IV.59a})$$

$$\text{var}(N_0) = (NF_e)^2 \left[\frac{\text{var}(N)}{N^2} + \frac{\text{var}(F_e)}{F_e^2} \right], \quad (\text{IV.59b})$$

and hence, unless N is very small,

$$\text{var}(N_0) \approx N_0^2 \left[\frac{\text{var}(N)}{N^2} + \frac{1 - e^{-\lambda(t-t_0)}}{N} \right]. \quad (\text{IV.60})$$

If instead of extrapolating numbers of atoms we adopt the more usual course of extrapolating counts per unit time A at time t back to counts per unit time A_0 at time t_0 , if the efficiency of our counter is η we have

$$A = \eta\lambda N, \quad (\text{IV.61a})$$

$$A_0 = \eta\lambda N_0, \quad (\text{IV.61b})$$

$$A_0 = AF_e, \quad (\text{IV.61c})$$

and consequently, for N considerably greater than 1,

$$\text{var}(A_0) = A_0^2 \left[\frac{\text{var}(A)}{A^2} + \frac{1 - e^{-\lambda(t-t_0)}}{A} \cdot \lambda\eta \right]. \quad (\text{IV.62})$$

The last term within the parenthesis implies that if we wish to know the variance of an extrapolated number, we must know the efficiency of our counter; we are faced with the necessity of absolute counting.

Fortunately, the nature of the expression permits a simplifying approximation which in most cases eliminates this requirement. The "extrapolation variance" has an upper limit; since η is at most 1, its asymptotic maximum value at long times is λ/A . If this value is small compared to $\text{var}(A)/A^2$ the extrapolation error may be neglected without seriously affecting our estimate of the precision of our answer.

7. Statistical Reliability of Measurements on Nuclear Reactions

Let us consider an experiment in which a sample of M atoms is immersed in a flux ϕ of particles per unit area per second and left there for a time t . The sample is then counted immediately and n events are recorded. It is desired to calculate the cross section σ of the reaction, and its variance, from the data.

This problem can most readily be approached by redefining p to be the probability that an atom of original target material will react and subsequently give rise to a count in the apparatus. From this definition it will readily be seen that, in order for p to be appreciably greater than zero, all of the following conditions must be met: (1) An appreciable fraction of the target material must be consumed in the nuclear reaction; (2) the reaction must be carried out in a time short compared to the product half-life; (3) the counting measurement must be carried out over a time long compared to the half-life; (4) a major portion of the sample must be placed in the counter; and (5) the counter must be of high geometry and high efficiency. These conditions are mutually exclusive under present technology, since conditions (1) and (2) require enormous fluxes available only in nuclear explosions and condition (4) is not possible to fulfill following a nuclear explosion. Consequently p is always small, and Poisson statistics always applies when deducing the variance of a nuclear reaction parameter such as flux, reaction cross section, or amount of target material.

V. THE RESOLUTION OF DECAY DATA

1. Nature of the Problem

We have seen in Chapter II that the expected value of the activity of a sample of radioactive material changes with time in a manner which may be described by a linear combination of negative exponential functions, that is, functions of the form $\exp(-\lambda_i t)$, where each λ_i is associated with one of the species present so that for n species present we have n terms, plus a possible constant term if some species are being formed via a nuclear reaction. We have seen further that certain physical properties, such as reactor flux or reaction cross section, may be derived from the values of the coefficients of these exponential terms. It remains to apply the statistical techniques of Chapters III and IV to the resolution of decay data, that is, to determining the best values for such coefficients from a set of experimental observations.

We have seen in Chapter II how these coefficients may be estimated graphically; however, for the best solution to our problem it is necessary to resort to a least-squares solution following methods described in Chapter III, Section 5. The data (obtained as events per unit time) are to be fitted to a function of the form

$$y_j = \sum_{i=1}^n x_i e^{-\lambda_i t_j} \quad (V.1)$$

where t_j is measured at the midpoint of each measured count, or for extreme precision (see Chapter II, Section 8)

$$y_j = \sum_{i=1}^n x_i e^{-\lambda_i t_j} \left[\frac{1 - e^{-\lambda_i (\Delta t)_j}}{\lambda_i (\Delta t)_j} \right] \quad (V.2)$$

where $(\Delta t)_j$ is the duration of the j th counting measurement and t_j is measured at the start of the count. The coefficients to be determined are the x_i ; proper weighting will be discussed below in detail.

We shall assume in this chapter that the λ are precisely known for all the species in the system. We shall consider the proper weighting of the data for a least-squares solution, the kind and magnitude of the variances which contribute to the weighting factor, and the effect on the results of any unanticipated variance. The λ for our solution are assumed to be constants; we

shall investigate the systematic effect on the resolution of slight variations in the values of the λ . We shall also consider briefly the problem of rejection of data which appear to fall outside the expected range of variation.

In addition, we shall also consider briefly the analysis of replicate samples, the analysis of the sources of variation between replicates, and the proper averaging of replicate results.

2. The Least-Squares Solution; Weighting of the Data

We have seen in Chapter III that the squares of the deviations of points y_j should be weighted by the reciprocals of the variances of the y_j before summing for the least-squares solution; alternatively, each equation describing a bit of data y_j should be weighted by the reciprocal of the square root of the variance of y_j before commencing a matrix solution.

If every count is recorded, and the counting equipment is perfectly stable, the variance of an observed count is given by equation IV.50b of Chapter IV. In the usual case, the probability of a single atom being counted is negligibly small and the gross number of events recorded is very large; the variance of the event is then the gross number of events, plus the number of background events observed in a measurement of equal duration. Frequently, the background is observed in a measurement of one duration t_B and the gross count is observed in another of different duration t_G ; if the number of gross and background events are G and B respectively, and the desired net counting rate is z , then

$$z = \frac{G}{t_G} - \frac{B}{t_B} \equiv g - b, \quad (\text{V.3})$$

and

$$\text{var}(z) = \text{var}(g) + \text{var}(b) \quad (\text{V.4})$$

by Theorem IV of Chapter III, since measurements of g and b are independent. By equation III.14,

$$\text{var}(z) = \text{var}\left(\frac{G}{t_G}\right) + \text{var}\left(\frac{B}{t_B}\right) = \frac{1}{t_G^2} \text{var}(G) + \frac{1}{t_B^2} \text{var}(B). \quad (\text{V.5})$$

When the Poisson distribution function is applicable,

$$\text{var}(z) = \frac{G+1}{t_G^2} + \frac{B+1}{t_B^2}, \quad (\text{V.6})$$

and when both G and B are large compared to unity,

$$\text{var}(z) = \frac{g}{t_G} + \frac{b}{t_B} . \quad (\text{V.7})$$

If a scaling circuit is used in which the register records only every S_G events during the gross measurement, and every S_B events during the background measurement, and if only integral numbers of registered events are recorded without regard to incompleting sets of S_G or S_B events, an additional source of variance is introduced. Two types of circuits are in common use: in the first type, any events arriving after the last completed register are ignored when the scaling circuit is reset and recorded, so that a measurement which is only 1 count below a completed set of $N + 1$ registered sets of events is recorded as N registered sets of events; in the second type, the reading on the register is increased by 1 if the next set of S counts has been at least half completed. With the first type, a reading of N registered sets of events implies a number of events in the counter anywhere between NS and $(N + 1) \times S - 1$; with the second type, a reading of N registered sets of events implies a number of events in the counter between $(N - \frac{1}{2})S$ and $(N + \frac{1}{2})S - 1$. Any value in the range is equally likely in either case. The expected value of the number of events in the counter is obviously $(N + \frac{1}{2})S - \frac{1}{2}$ in the first case and $NS - \frac{1}{2}$ in the second case; the distribution function in each case is seen to be a rectangular function $S - 1$ counts wide. For normalization, the height of the rectangle must be $\frac{1}{S - 1}$, so that in the first case

$$P_x(u) = \begin{cases} 0, & u < NS, \\ \frac{1}{S - 1}, & NS \leq u \leq (N + 1)S - 1, \\ 0, & u > (N + 1)S - 1, \end{cases} \quad (\text{V.8a})$$

and in the second case,

$$P_x(u) = \begin{cases} 0, & u < (N - \frac{1}{2})S, \\ \frac{1}{S - 1}, & (N - \frac{1}{2})S \leq u \leq (N + \frac{1}{2})S - 1, \\ 0, & u > (N + \frac{1}{2})S - 1. \end{cases} \quad (\text{V.8b})$$

The variance of this distribution may be obtained from equation III.33 for either case,

$$\text{var}(N) = \int_{\left(N-\frac{1}{2}\right)S}^{\left(N+\frac{1}{2}\right)S-1} \left(x - NS + \frac{1}{2}\right)^2 \frac{dx}{(S-1)} = \int_{NS}^{(N+1)S-1} \times \left(y - \left(N + \frac{1}{2}\right)S + \frac{1}{2}\right)^2 \frac{dy}{(S-1)}. \quad (\text{V.9})$$

The identity of the two integrals may be seen by a transformation of variables, $y = x + \frac{1}{2}S$. Upon performing the integration and simplifying, the result is

$$\text{var}(n) = \frac{(S-1)^2}{12}. \quad (\text{V.10})$$

This variance must be added to each of the variances included in equation V.5 for both gross and background measurements, so that in place of equation V.7 we have

$$\text{var}(z) = \frac{g}{t_G} + \frac{b}{t_B} + \frac{(S_G - 1)^2}{12t_G^2} + \frac{(S_B - 1)^2}{12t_B^2}. \quad (\text{V.11})$$

When high scaling factors are used, the variance arising from neglect of the "interpolation lights" between even registers can affect the computed variance quite appreciably.

A final term to be included in the estimate of the variance of each measurement depends on the counting equipment. It seems reasonable to assume that the equipment itself is subject to variation, for example in the positioning of the sample on consecutive measurements or in the level of the counter discriminator; it also seems reasonable to assume that the effect of such variation on the counting rate is proportional to the counting rate. We therefore postulate an additive variance for both gross and background counting rates, proportional to the square of the counting rate; the constant of proportionality will be dependent on the actual instrument and perhaps on the species counted.

Assuming for the moment that the constant of proportionality depends only on the instrument, we must realize that the sample positioning mechanism is part of the instrument for the gross count, but not for the background count; the constants for the two measurements will therefore be different. Estimation of such constants is extremely difficult, especially when they are small, since the only measurement that can be made also involves random variances of the types described above; a technique will be discussed later.

If the variance due to counter variation is $\gamma^2 g^2$ for the gross count and $\beta^2 b^2$ for the background count rate, equation V.11 then becomes

$$\text{var}(z) = \frac{g}{t_G} + \frac{b}{t_B} + \frac{(S_G - 1)^2}{12t_B^2} + \frac{(S_B - 1)^2}{12t_B^2} + \gamma^2 g^2 + \beta^2 b^2. \quad (\text{V.12})$$

This expression is used to evaluate the proper weighting factors for use in a least-squares calculation of decay data, as described in Chapter IV.

A matrix inversion method is highly desirable, since the inverse matrix of the normal equations (see equations III.92d, 104f, 130, and 142g)(multiplied by $\frac{\chi^2}{m-n}$ when this term is greater than unity) is the covariance matrix of the parameters; consequently, the diagonal elements of that matrix are the variances of the x_i . We then have an estimate of the precision of the amount of each component present, since the standard error of x_i is the square root of its variance.

If the variance due to counter variation is appreciable and is also dependent on the species being counted, proper weighting for a least-squares solution for a multicomponent system becomes very laborious. Estimation of the variance of a given measurement a priori is impossible, since the proportions of the various kinds of radiation making up the observed count are not known until after the resolution is done; if the counter variation constants for the various species are known from independent measurements on single-component systems, the calculation may be done by successive approximations. Such a case may arise when the amplifying circuit of the counter has an energy cutoff which is subject to random drift; radiations producing pulses in the counter close to the cutoff level will be much more severely affected than those which produce pulses well removed from the cutoff. The practice of counting single photopeaks for decay, discriminating out pulses of both higher and lower energy than the gamma radiation of interest, is particularly subject to variation of this kind.

3. Estimation of Variance Due to Instrumental Variation

In order to make an estimate of the variance due to instrumental variation from a set of experimental data, it is necessary to resort to an iterative

process. If the variance due to instrumental background is desired it can be obtained from a series of background measurements by the technique to be described below; in almost all cases it will be trivially small compared to the instrumental variation while counting an actual sample. In what follows we neglect it.

We resort to the χ^2 test described in Chapter III, Section 5 (equations III.138, 139, and 140) to give us an indication of the magnitude of instrumental variance. We perform a least-squares solution using equation V.12 to give the weighting factors, assuming initially that the constant γ is zero. If χ^2 is significantly larger than its expected value, we repeat the solution using a trial value of γ which is believed to be reasonable. The new value of χ^2 will of necessity be smaller; if it is still too large, the trial value is increased and the calculation repeated again. Once a trial value of γ is found which reduces χ^2 below its expected value, a "best" value of γ can be approached with a limited number of iterations by any one of a number of schemes of interpolation between trial values.

This system makes the obviously false assumption that the observed value of χ^2 for this particular set of observations is precisely equal to its expected value; the "best value" of γ found from any one set of data is relatively worthless. However, it seems reasonable to assume that the observed χ^2 values of a number of sets of good data, properly weighted, will be distributed about the expected value; consequently we expect that the set of calculated "best values" of γ will be distributed about the true value. A large number of sets of observations should then give a number of values of γ which, when averaged, should give a very reasonable estimate of γ .

Such a procedure involves a very large amount of sample preparation, counting, and especially calculation; if the experimental data to be processed are not of the highest precision the effort is probably not justified. The use of high speed computers for the iterative calculations is practically mandatory.

A more rapidly converging method can be derived as follows: Let

Y_i = the "true" value of the count rate at time t_i ,

η_i = the least-squares computed value of the rate at t_i ,

y_i = the observed value of the rate at t_i ,

r_i = the residual at t_i ,

that is,

$$r_i = \eta_i - y_i, \quad (V.13)$$

where η_i is computed without allowance made for instrumental variance.

Now to a good approximation,

$$E(r_i^2) \approx E\left[\left(Y_i - y_i\right)^2\right] = \sigma_i^2 + \gamma^2 y_i^2 \quad (V.14)$$

so that

$$E\left[\frac{r_i^2 - \sigma_i^2}{y_i^2}\right] \approx \gamma^2. \quad (V.15)$$

If there are m observations and n components there are $m - n$ degrees of freedom in the system. Therefore we expect that

$$\frac{1}{m - n} \sum_i \left(\frac{r_i^2 - \sigma_i^2}{y_i^2}\right) \approx \gamma^2. \quad (V.16)$$

In three cases selected at random an estimation of γ by this method gave values which were within 2% (in each case) of the value obtained by the iteration method after a dozen iterations. The value so obtained can obviously be used to recompute the η_i to permit an iterative improvement in γ .

4. Effect of Input Parameters on Resolution

We have considered the decay constants λ_i of the system under consideration to be fixed and absolutely known up to this point. A detailed consideration of the determination and precision of the λ_i must be deferred to the following chapter; it is, however, convenient to consider at this point the effect of a slight change in one of the λ_i on the derived values of the system parameters x_i and on the goodness of fit of the data to the calculated curve, as described by χ^2 .

The set of n normal equations for the system as described in Chapter III, equation III. 92d with appropriate substitution, are

$$\sum_{i=1}^n \left[\sum_{j=1}^m \frac{e^{-(\lambda_i + \lambda_k)t_j}}{\text{var}(y_j)} \right] x_i = \sum_{j=1}^m \frac{y_j e^{-\lambda_k t_j}}{\text{var}(y_j)} \quad (V.17)$$

for a set of m observations on n components, where k has each value from 1 to n . We desire to know the effect on each x_i of a small variation in a particular λ , symbolized by λ_s ; in other words, we want the quantity

$\frac{\partial x_i}{\partial \lambda_s}$ for each i and for each s .

If one performs the operation of partial differentiation with respect to λ_s on the set of equations V.17, one obtains the set of equations ($1 \leq k \leq n$)

$$\sum_{i=1}^n \left[\sum_{j=1}^m \frac{e^{-(\lambda_i + \lambda_k)t_j}}{\text{var}(y_j)} \right] \frac{\partial x_i}{\partial \lambda_s} - \sum_{i=1}^n \left[x_i \sum_{j=1}^m \frac{t_j (\delta_{is} + \delta_{ks}) e^{-(\lambda_i + \lambda_k)t_j}}{\text{var}(y_j)} \right] = - \sum_{j=1}^m \frac{t_j \delta_{ks} y_j e^{-\lambda_k t_j}}{\text{var}(y_j)} \quad (\text{V.18})$$

where

$$\begin{aligned} \delta_{ij} &= 0, \quad i \neq j \\ &= 1, \quad i = j, \end{aligned}$$

which upon rearrangement gives

$$\sum_{i=1}^n \left[\sum_{j=1}^m \frac{e^{-(\lambda_i + \lambda_k)t_j}}{\text{var}(y_j)} \right] \frac{\partial x_i}{\partial \lambda_s} = \sum_{i=1}^n \left[x_i (\delta_{is} + \delta_{ks}) \sum_{j=1}^m \frac{t_j e^{-(\lambda_i + \lambda_k)t_j}}{\text{var}(y_j)} \right] - \delta_{ks} \sum_{j=1}^m \frac{y_j t_j e^{-\lambda_k t_j}}{\text{var}(y_j)} \quad (\text{V.19})$$

for any value of $1 \leq s \leq n$. If this set of equations is expressed in matrix form it will be seen at once that the matrix of the coefficients of the new set of unknowns $\partial x_i / \partial \lambda_s$ is identical with the matrix of the normal equations; if these have already been solved by matrix inversion, the inverse matrix is already available. The values of the x_i from the solution of the normal equations may be inserted into the right members of the set of equations to give the components of a vector v_s which, multiplied from the left by the inverse matrix of the normal equations, yields the vector whose components are $\partial x_i / \partial \lambda_s$. If a matrix P is formed whose columns are the set of vectors v_s with s taking on the values 1 to n , in turn, then this matrix multiplied

from the left by the inverse matrix of the normal equations will give a new matrix Q such that

$$Q_{is} = \frac{\partial x_i}{\partial \lambda_s} \quad (V.20)$$

which is a compact solution to our problem.

It is perhaps more convenient to express the components of Q in terms of fractional changes and to convert Q to a matrix Q' such that

$$Q'_{is} = -\frac{\lambda_s}{x_i} \frac{\partial x_i}{\partial \lambda_s} = \frac{T_s}{x_i} \frac{\partial x_i}{\partial T_s} \quad (V.21)$$

where

$$T_s = \frac{\ln 2}{\lambda_s} .$$

The negative sign is inserted to put the significance of the number so derived in terms of an alteration in the half-life value, rather than the decay constant, since most workers are used to thinking in terms of half-lives rather than decay constants.

We can now derive the effect of varying an input λ on the fit of the data to the calculated curve. We have

$$\chi^2 = \sum_{j=1}^m \frac{r_j^2}{\text{var}(y_j)} = \sum_{j=1}^m \frac{\left(y_j - \sum_{i=1}^n x_i e^{-\lambda_i t_j} \right)^2}{\text{var}(y_j)} . \quad (V.22)$$

Therefore

$$\frac{\partial(\chi^2)}{\partial \lambda_s} = \sum_{j=1}^m \frac{2 \left(y_j - \sum_{i=1}^n x_i e^{-\lambda_i t_j} \right) \left[t_j x_s e^{-\lambda_s t_j} - \sum_{i=1}^n e^{-\lambda_i t_j} \left(\frac{\partial x_i}{\partial \lambda_s} \right) \right]}{\text{var}(y_j)} . \quad (V.23)$$

From the least-squares solution we get the values of the x_i and from the matrix Q described above we get the values of the $\partial x_i / \partial \lambda_s$. A positive value of this derivative means λ_s should be reduced (i. e., the half-life should be

increased) to improve the fit; a negative value means the reverse. This calculation can be made the basis of an iterative method for improving the values of the input λ , providing that fairly good values are available initially.

5. Criteria for the Rejection of Data

It frequently happens that a set of data contains one or more points which do not actually belong to the set, due either to an instrument malfunction or to human error on the part of the experimenter. Typical errors are: misreading the instrument; counting the wrong sample; counting the sample in the wrong counting position; failing to remove or insert an absorber; failing to remove a previous sample; failing to set proper discriminator levels; failing to record the data in legible form, resulting in misreading of the written data; and so on ad nauseam. If the instrument (or the operator) is so erratic that many points are subject to such nonstatistical errors, the experiment is worthless; there is no way of selecting the good data from the total set a priori. If no further information is available, the experiment should be redesigned. If the fault is in the operator, suitable corrective action sometimes requires considerable ingenuity.

If the set of data contains only one or two erroneous points, these can usually be detected by inspection of a graphical presentation of sample activity plotted against time. This is, however, a subjective method and takes no account of the precision of the individual measurements. An objective criterion for rejection of suspected data is desirable.

For convenience in evaluating data, it is highly desirable to tabulate the points in order of time, with their residuals expressed both as "percentage error" and as "number of standard deviations," i. e., the weighted residual. Whenever the magnitude of the weighted residual exceeds about 3, the point should be examined further, since such a large deviation is highly unlikely to occur through random error. If the weighted residual is unreasonably large, but the percentage error is quite small (as often happens when a very large number of events is recorded) then no appreciable improvement in the answer is achieved by rejecting the point in question.

One is then led to the idea that a point should be rejected if: (1) The point is separated from the "true" calculated curve by more than about three of its standard deviations; and (2) Omission of the point in question alters

the final answer appreciably. The "true" calculated curve may be estimated from the curve calculated without the suspected point. A systematic check of every point in a set of m points, then, will require a series of $m + 1$ least-squares calculations - a very laborious task.

It is possible to approximate the same criterion by some far simpler tests. One can estimate that a human error will cause a percentage deviation of the observed point greater than a certain arbitrarily selected level, say 5%; or alternatively one can decide that rejection of a single point in the set which is within 5% of the computed curve (using all points) will not affect the answer appreciably. One can then class all points with a percentage deviation from the calculated curve of 5% or greater, and a weighted residual of 3 or greater, as "suspect."

Unfortunately, a simple scheme which rejects all such points, or even rejects them one at a time starting with the point with the largest deviation, proves in practice to have serious flaws - in fact, it retains bad points and rejects good ones a fair fraction of the time. An additional criterion which has been found to give good results in our laboratory is based on the following approach: assume the existence of a set of "good" data points with a calculated curve passing through them. The deviations of the points from the curve will be distributed, some positive and some negative, none very large, and more or less at random. If now a "bad" point is added to the data and the curve is recomputed, the presence of the bad point will tend to displace the new computed curve toward that point and away from the good points on either side. The new curve will then pass between the good points and the bad one in the region of the bad point. This gives us our additional criterion: for automatic rejection, the deviation of the suspect point (suspect in the sense described above) from the calculated curve must be in the opposite sense from the deviations of the points on either side of it; or if the suspect point is the first or last point, its deviation must be in the opposite sense from that of the next point to it.

This triple criterion has been in use in Livermore for some time and has been applied to literally hundreds of sets of data. It has occasionally caused unneeded rejection of a good point, usually when the bad point is the second data point - in this case the first point is sometimes rejected also - but if data are sufficiently redundant this usually does no harm.

If there are suspect points in the data, and if the third criterion for rejection is not met, one should examine the fit of the data to the calculated curve. Often a systematic pattern will show up in the residuals, most easily seen as a grouping of positive and negative residuals in clusters. This grouping indicates an improper form for the function being fitted - the sample is impure, or the wrong half-lives are being used.

If the analysis of the data is performed by a programmed computer, it is very desirable to obtain an output which presents the input parameters, the results, a tabulation of points and their deviations, and the results of the χ^2 test. A common complaint of those using least-squares analysis in an improper manner is that any function can be fitted to any set of data by a least-squares method; while this is undoubtedly true, it is easy to avoid gross errors if one considers not only the answers so obtained but also their precision, the goodness of fit as given by χ^2 , and the pattern shown by the residuals. Also, providing that suitable computing facilities are available, a great deal of labor may be saved and a great many experiments protected to some degree against subjective bias.

6. Avoidance of Computer Overflow

It occasionally happens that a multicomponent system has in it a short-lived daughter of a longer-lived parent, and that the purification of the parent species is performed some time after the time of interest. If the activity measurement is carried out in such a way that the daughter component must be included in the resolution, i. e., if the growth of the daughter is observed, difficulties may arise in the analysis if precautions are not taken. The time elapsed between zero-time and the time of measurement may be, and often is, so large with respect to the daughter half-life that an attempt to evaluate the necessary exponential functions will result in a computer overflow. This problem can always be avoided by a substitution of a spurious zero-time, conveniently taken as the time of the first count; the resolution can be performed to give results as of that time and these results extrapolated back to the true zero-time. No difference will be obtained in the answers by this method. The advantage gained is that a separate exponential extrapolation factor is used for each component; the exponent can be evaluated and tested for magnitude before exponentiation, and one can refrain from extrapolating those components whose exponents are too large. Such coefficients are without physical meaning in any case.

Since extrapolation involves merely multiplying by a constant, the fractional precisions of the x_i are not changed by this process. However, if one is interested in the sensitivity of the x_i to the input half-lives, the Q-matrix (see equation V-20) as derived by this method will be changed by extrapolation. Let

$$y_i = x_i e^{\lambda_i \Delta t} \quad (V.24)$$

where Δt is the time between the true zero-time and the time of the first count. Then

$$\frac{\partial y_i}{\partial \lambda_j} = \frac{\partial x_i}{\partial \lambda_j} e^{\lambda_i \Delta t} + \delta_{ij}(\Delta t) x_i e^{\lambda_i \Delta t} \quad (V.25)$$

where δ_{ij} has the usual meaning.

7. Analysis of Replicate Samples; Analysis of Variance

It is the usual practice to do replicate analyses on radiochemical samples whenever possible. A set of replicates will, in general, give a set of approximately equal answers; since the result of a radiochemical analysis is subject to statistical imprecision, we would not expect replicate samples to give identical answers except fortuitously. We must bear in mind also that the statistical imprecision of a single radiochemical analysis, as estimated from the least-squares solution, does not give a true picture of the precision of the result; there are analytical variations possible which have a systematic, rather than a random, effect on all measurements performed on a single sample. Examples are: an error in the measurement of the aliquot taken for analysis; an error in the measurement of the recovered yield; an inhomogeneity in the mounted counting sample, resulting in a systematic perturbation in the counting rate; an error in the amount of carrier added; and so on. We assume that these analytical variations have a random effect on the answers from a set of replicate samples. It remains to develop a technique for estimating the imprecision due to such variations from the available data. It should also be noted that there are numerous systematic errors which can distort the results of an entire set of replicates. Elimination of such errors is a matter of careful experimental technique; no amount of data manipulation can either detect or compensate for them.

Each replicate answer may be considered to be the expected value of a probability distribution function of known variance. These expected values

do not all coincide. Furthermore, since these distribution functions are not all of equal variance, their worths in establishing an overall "best" value are not all equal. We desire to analyze the results of such replicate analyses mathematically in order to make an estimate of the imprecision due to analytical technique. When this is available, one can then add the variance due to analytical technique to the variance due to statistical fluctuations in the decay data to get the total variance for each replicate; the reciprocal of this total variance is then used as a weighting factor to average the replicate results, and their variances, as before to give a best estimate of the true value for the experiment and its variance.

The problem is analogous to one discussed by Davies⁹ in which a number of analyses n_i is performed on each of k samples, the variance of an individual analysis is σ_0^2 and the variance between samples is σ_1^2 . The numbers n_i are not necessarily equal between samples. It can be shown that, if x_{ij} represents the result from the j th replicate analysis on the i th sample, the quantity

$$M_1 = \frac{1}{k-1} \left\{ \sum_{i=1}^k \frac{\left(\sum_{j=1}^{n_i} x_{ij} \right)^2}{n_i} - \frac{\left(\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij} \right)^2}{\sum_{i=1}^k n_i} \right\} \quad (V.26)$$

is an estimate of $\sigma_0^2 + \bar{n}\sigma_1^2$, where

$$\bar{n} = \frac{\left(\sum_{i=1}^k n_i \right)^2 - \sum_{i=1}^k n_i^2}{(k-1) \sum_{i=1}^k n_i} \quad (V.27)$$

Consequently, an estimate of σ_1^2 is given by

$$\sigma_1^2 = \frac{M_1 - \sigma_0^2}{\bar{n}} \quad (V.28)$$

⁹Owen L. Davies, Editor, Statistical Methods in Research and Production (Hafner Publishing Co., N. Y., 1958), III Ed. (revised), pp. 134-136.

In order to apply this result to our problem, we consider our various results x_i to be averages of undetermined numbers n_i of analyses each with a variance σ_0^2 ; from Chapter III, equation III.51, $\text{var}(x_i)$ is given by

$$\text{var}(x_i) = \frac{1}{n_i} \cdot \sigma_0^2 \quad (\text{V.29})$$

and consequently

$$n_i = \sigma_0^2 \cdot \frac{1}{\text{var}(x_i)} \quad (\text{V.30})$$

Since we have assumed that

$$x_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} \quad (\text{V.31})$$

then

$$\sum_{j=1}^{n_i} x_{ij} = n_i x_i = \frac{x_i}{\text{var}(x_i)} \cdot \sigma_0^2 \quad (\text{V.32})$$

Consequently,

$$M_1 = \frac{\sigma_0^2}{k-1} \left\{ \sum_{i=1}^k \frac{x_i^2}{\text{var}(x_i)} - \frac{\left[\sum_{i=1}^k \frac{x_i}{\text{var}(x_i)} \right]^2}{\sum_{i=1}^k \frac{1}{\text{var}(x_i)}} \right\} \quad (\text{V.33})$$

and

$$\bar{n} = \frac{\sigma_0^2}{k-1} \left\{ \frac{\left[\sum_{i=1}^k \frac{1}{\text{var}(x_i)} \right]^2 - \sum_{i=1}^k \frac{1}{[\text{var}(x_i)]^2}}{\sum_{i=1}^k \frac{1}{\text{var}(x_i)}} \right\}, \quad (\text{V.34})$$

and therefore, substituting in V.28 and simplifying,

$$\sigma_1^2 = \frac{\left[\sum_{i=1}^k \frac{1}{\text{var}(x_i)} \right] \left[\sum_{i=1}^k \frac{x_i^2}{\text{var}(x_i)} \right] - \left[\sum_{i=1}^k \frac{x_i}{\text{var}(x_i)} \right]^2 - (k-1) \sum_{i=1}^k \frac{1}{\text{var}(x_i)}}{\left[\sum_{i=1}^k \frac{1}{\text{var}(x_i)} \right]^2 - \sum_{i=1}^k \frac{1}{[\text{var}(x_i)]^2}} \quad (\text{V.35})$$

Note that $\text{var}(x_i)$ as used above refers only to the variance of the sample due to fluctuations in the decay data; in order to average the replicate values as was done in Chapter III, equations III.77 and 78, the estimate for σ_1^2 should be added to each variance, so that for the best average value of k replicates,

$$\bar{x} = \frac{\sum_{i=1}^k \frac{x_i}{\text{var}(x_i) + \sigma_1^2}}{\sum_{i=1}^k \frac{1}{\text{var}(x_i) + \sigma_1^2}} \quad (\text{V.36})$$

and

$$\text{var}(\bar{x}) = \frac{1}{\sum_{i=1}^k \frac{1}{\text{var}(x_i) + \sigma_1^2}} \quad (\text{V.37})$$

A derivation of equation V.35 follows:

Let the results of our replicate analyses be given by x_i , with each of which is associated a variance $\text{var}(x_i)$ derived exclusively from counting statistics. $\text{var}(x_i)$ is then obtained from the least-squares resolution of the decay curve of the i th sample. We assume that each x_i is equivalent to an average of n_i virtual individual results, each with the same variance σ_0^2 .

Given any pair of virtual measurements, x_{ij} and $x_{\ell m}$, let their mean be represented by $\mu_{ij\ell m}$; then an estimate of the variances σ_0^2 and σ_1^2 may be obtained from this pair of measurements as follows:

If $i = \ell$, then σ_1^2 is not estimated, since the two samples are from the same group.

If $i = \ell$ and $j = m$, no variance is estimated.

Let the variance estimated from our pair be

$$V_{ijlm} = \frac{(x_{ij} - \mu_{ijlm})^2 + (x_{lm} - \mu_{ijlm})^2}{2 - 1}, \quad (V.38)$$

and

$$V_{ijlm} \approx \sigma_0^2 (1 - \delta_{il} - \delta_{jm}) + \sigma_1^2 (1 - \delta_{ij}). \quad (V.39)$$

Substituting in V.39 for V_{ijlm} , given that

$$\mu_{ijlm} = \frac{x_{ij} + x_{lm}}{2} \quad (V.40)$$

we get

$$\frac{1}{2} (x_{ij} - x_{lm})^2 \approx (1 - \delta_{ij} \delta_{lm}) \sigma_0^2 + (1 - \delta_{ij}) \sigma_1^2. \quad (V.41)$$

Summing over all possible pairs to get the best estimate, we get

$$\begin{aligned} \sum_{i=1}^k \sum_{j=1}^{n_i} \sum_{\ell=1}^k \sum_{m=1}^{n_\ell} \left(\frac{x_{ij}^2 + x_{lm}^2}{2} - x_{ij} x_{lm} \right) &\approx \sigma_0^2 \sum_i^k \sum_j^{n_i} \sum_\ell^k \sum_m^{n_\ell} (1 - \delta_{ij} \delta_{lm}) \\ &+ \sigma_1^2 \sum_i^k \sum_j^{n_i} \sum_\ell^k \sum_m^{n_\ell} (1 - \delta_{ij}), \end{aligned} \quad (V.42)$$

which, after a little algebra, and letting $N = \sum_i^k n_i = \sum_\ell^k n_\ell$ for ease of writing,

gives

$$N \sum_i^k \sum_j^{n_i} (x_{ij}^2) - \left(\sum_i^k \sum_j^{n_i} x_{ij} \right)^2 \approx \sigma_0^2 (N^2 - N) + \sigma_1^2 \left(N^2 - \sum_{i=1}^k n_i^2 \right). \quad (V.43)$$

Now σ_0^2 may be estimated within each sample by

$$\sigma_0^2 \approx \frac{\sum_{j=1}^{n_i} (x_{ij} - x_i)^2}{n_i - 1}, \quad (V.44)$$

so that

$$\sum_{j=1}^{n_i} (x_{ij} - x_i)^2 \approx (n_i - 1)\sigma_0^2. \quad (V.45)$$

Again summing over all samples to obtain the best estimate, we get

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij}^2) - \sum_{i=1}^k n_i x_i^2 \approx (N - k)\sigma_0^2 \quad (V.46)$$

which gives

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij}^2) \approx \sum_{i=1}^k n_i x_i^2 + (N - k)\sigma_0^2. \quad (V.47)$$

Substituting this result into V.43 and rearranging, we get

$$\sigma_1^2 = \frac{N \sum_{i=1}^k n_i x_i^2 - \left(\sum_{i=1}^k n_i x_i \right)^2 - N(k - 1)\sigma_0^2}{N^2 - \sum_{i=1}^k n_i^2}. \quad (V.48)$$

If we now insert our definition of N and substitute for n_i from equation V.30, upon simplification we get the desired result, equation V.35. It will be observed that the value of σ_0^2 is indeterminate, as it must be since it represents the variance of a distribution of virtual objects postulated for convenience.

A negative value of σ_1^2 signifies that, fortuitously, the replicate measurements agree better than one would expect from their precision. This occurrence is analogous to a value of count variance less than its expected value, as discussed on page 56; such a result merely implies that there is no clear evidence for analytical variation, and therefore σ_1^2 should be given the value zero in Equations V.36 and V.37. Frequent appearance of negative values of σ_1^2 in a set of results indicates that errors of the input numbers are being improperly assigned.

VI. THE IDENTIFICATION OF COMPONENTS AND MEASUREMENT OF HALF-LIVES OR DECAY CONSTANTS

1. Nature of the Problem

We have been operating on the assumption that the decay constants of the components of our systems have been precisely known. While it is true that the decay constant of a nuclear species is not a variate, in the sense that it is not subject to variation due to sampling, nevertheless it must be realized that the only method of evaluating a decay constant is by observing the process of decay - which inevitably involves random statistical errors. In this sense, then, a decay constant must be treated as a variate, although its value must be considered fixed during any one resolution calculation - variation in λ makes a systematic, rather than a random, variation in the results of a resolution. In this chapter we shall consider several different approaches to the measurement of λ or its equivalent, the measurement of half-life, from a set of data, together with the inseparable problem of the detection of the number of components in the system.

It is of the utmost importance in measuring the half-life of a species to make sure that the sample under examination is free of even the smallest amount of radioactive impurities. A grossly impure sample is often easy to identify; if the species involved are reasonably short-lived, the decay curve will clearly show the presence of more than one component unless the half-life of the impurity is similar to that of the species of interest. If, however, the extent of contamination is slight, or the species involved are too long-lived to decay to any great extent in a reasonable period, detection of an impurity is more difficult. It can sometimes be accomplished by analysis of the radiations from the sample; sometimes the only recourse is to observe the specific activity of the sample as it is subjected to more and more rigorous chemical purification. If the impurity is isotopic, it is necessary either to use an isotope separator to prepare a pure sample, or to estimate the desired half-life by a treatment of the data which takes account of the presence of components other than the one of interest.

An example of the effect of an impurity on the measurement of a half-life can be found in the early measurements of the half-life of Cd^{115} formed from fission. The original reports on this nuclide assigned it a half-life of

58 hours; when the existence of the isomeric state, with a half-life of 43 days, was discovered, re-examination of the data resulted in a change to the current best value for Cd¹¹⁵ of 53 hours - a change of nearly 10%.

2. Single-Component System of Short or Moderate Half-Life

For simplicity, let us first consider a system which we are sure contains only one component during the entire period covered by the data. The decay of this system is given by

$$n = n_0 e^{-\lambda t} \quad (\text{VI.1})$$

which can be transformed into linear form by taking logarithms:

$$\ln n = \ln n_0 - \lambda t. \quad (\text{VI.2})$$

If n is the observed counting rate, we may evaluate λ graphically by plotting $\ln n$ against t and measuring the slope of the line; if semilogarithmic paper is used, we may measure half-life as the time interval required for n (not $\ln n$) to decrease by a factor of 2.

A more precise and objective result is desirable. If a redundant amount of data is available (i. e., more than two points) a least-squares solution can be performed as described in Chapter III, Section 5, using the set of equations VI.2 (one for each measured point). In order to weight the points properly we need the variance of $\ln n$; if the data are fairly precise this may be obtained approximately as follows:

$$\text{var} (\ln n) = \text{var} \left\{ \ln [E(n) + n - E(n)] \right\}, \quad (\text{VI.3a})$$

$$\text{var} (\ln n) = \text{var} \left(\ln \left\{ E(n) \left[1 + \frac{n - E(n)}{E(n)} \right] \right\} \right), \quad (\text{VI.3b})$$

$$\text{var} (\ln n) = \text{var} \left\{ \ln E(n) + \ln \left[1 + \frac{n - E(n)}{E(n)} \right] \right\}, \quad (\text{VI.3c})$$

$$\text{var} (\ln n) = \text{var} \left\{ \ln \left[1 + \frac{n - E(n)}{E(n)} \right] \right\}. \quad (\text{VI.3d})$$

Since $\frac{n - E(n)}{E(n)}$ is assumed to be small, we may expand the logarithm as a series and neglect all powers of $\frac{n - E(n)}{E(n)}$ above the first, so that

$$\text{var} (\ln n) \approx \text{var} \left[\frac{n - E(n)}{E(n)} \right], \quad (\text{VI.3e})$$

$$\text{var} (\ln n) \approx \text{var} \left[\frac{n}{E(n)} - 1 \right] = \text{var} \left[\frac{n}{E(n)} \right], \quad (\text{VI.3f})$$

$$\text{var} (\ln n) \approx \frac{\text{var}(n)}{[E(n)]^2}. \quad (\text{VI.4})$$

We then have the proper weighting factors to permit a least-squares analysis of the data for the parameters $\ln n_0$ and λ . If a matrix solution is used we have an objective evaluation of $\text{var} (\lambda)$, as described in Chapter III, equations III.92d, 104f, 130, and 142g; it is the proper diagonal element of the inverse matrix, multiplied by $\chi^2/(m-2)$ where this factor is greater than unity.

3. Single-Component System with Long Half-Life

If the half-life of the species under examination is so long that an experiment to observe the decay of a sample is difficult, another method must be used. There are three methods in general use. The simplest method depends on a direct measurement of specific activity; the second and more elaborate method depends on a careful comparison of the decay of the unknown species with that of a species of precisely known half-life; the third method depends on an internally calibrated method of observing decay over long periods of time.

Measurement of half-life by specific activity depends on the differential form of the fundamental law of radioactivity (equation II.1) which is

$$\frac{dn}{dt} = -\lambda n \quad (\text{VI.5})$$

where n is the number of atoms present and $-dn/dt$ is the disintegration rate. Direct application of this formulation of the law requires two difficult measurements: an absolute measurement of the disintegration rate of the sample, corrected for solid angle, counter efficiency, and sample absorption and scattering effects; and a measurement of the number of atoms in the sample. With the improvement in absolute counting techniques, and the increasing sensitivity of analytical instrumentation such as the mass spectrometer for measurement of n , this method is becoming more precise, but still requires great care. An example of the method, which gives some indications

of the care which must be used, will be found in Fleming, Ghiorso, and Cunningham's article¹⁰ on the half-life of uranium. A similar measurement was carried out at Harwell¹¹ to give the half-life of Cs¹³⁷.

A technique which has been used by several workers^{11,12} depends on the fact that two ionization chambers carefully made to be mechanically as nearly identical as possible will have nearly identical responses. A sample whose half-life is to be determined is placed in one chamber, and a very long-lived standard sample in the other. The two chambers are then connected to a sensitive current measuring device, e. g., an electrometer tube, in such a way that the difference in current between the two chambers may be measured. This difference is balanced to zero at the time t_0 and its change with time is noted. It will be seen that the "difference current" I_d will have the form

$$I_d = I_0 \left[1 - e^{-\lambda(t-t_0)} \right]. \quad (\text{VI.6})$$

If the experimental sample is removed from its chamber, the "difference current" observed with the standard sample compared to an empty chamber will be a measure of I_0 . We then have

$$\ln \left(1 - \frac{I_d}{I_0} \right) = -\lambda(t - t_0). \quad (\text{VI.7})$$

In practice, in particular if the decay of the sample is appreciably rapid, the exact time t_0 may be inconvenient to obtain; one may equally well select an arbitrary time t_a (near the time at which the chambers are exactly balanced, for convenience). Then let

$$X = \ln \left(1 - \frac{I_d}{I_0} \right) = -\lambda(t - t_a + t_a - t_0), \quad (\text{VI.8a})$$

$$X = \lambda(t_a - t_0) - \lambda(t - t_a). \quad (\text{VI.8b})$$

¹⁰ E. H. Fleming, Jr., A. Ghiorso, and B. B. Cunningham, *Phys. Rev.* 88, 642 (1952).

¹¹ F. Brown, G. R. Hall, and D. N. Walter, *J. Inorg. Nucl. Chem.* 1, 241 (1955).

¹² E. Segrè and C. E. Wiegand, *Phys. Rev.* 75, 39 (1949).

We may plot X against $(t - t_a)$ and obtain λ from the slope of the line, or for greater precision evaluate that slope by a least-squares method. Tobailem¹³ obtained the half-life of Co^{60} (5.27 years) to 1.3% in a series of measurements lasting only 32 days, and the half-life of Ac^{227} (21.6 years) to 2% in a series of measurements lasting only 137 days.

If the standard sample decays appreciably during the course of the measurement, and if its half-life is precisely known, then one inserts $I_0 e^{-\lambda_s(t-t_0)}$ into equation VI.7 in place of I_0 . In this case it is necessary to have a good estimate for t_0 ; this necessity becomes less urgent as the half-life of the standard becomes larger.

A self-calibrating method for measuring long-lived decays has been applied to the measurement of the half-life of Kr^{85} (ref. 14) and of Cs^{137} (ref. 15). The change in the ratio of $\text{Kr}^{85}/\text{Kr}^{86}$ in a gas sample, or $\text{Cs}^{137}/\text{Cs}^{135}$ in a fission-product cesium sample, can be followed for long times, since the fact that a ratio is measured is insurance against instrumental drift.

4. Multicomponent Systems - Detection

Given a sample or set of data about which no information (other than statistical reliability) is available, the first question which comes to mind is; is the system a single-component system, or is it not?

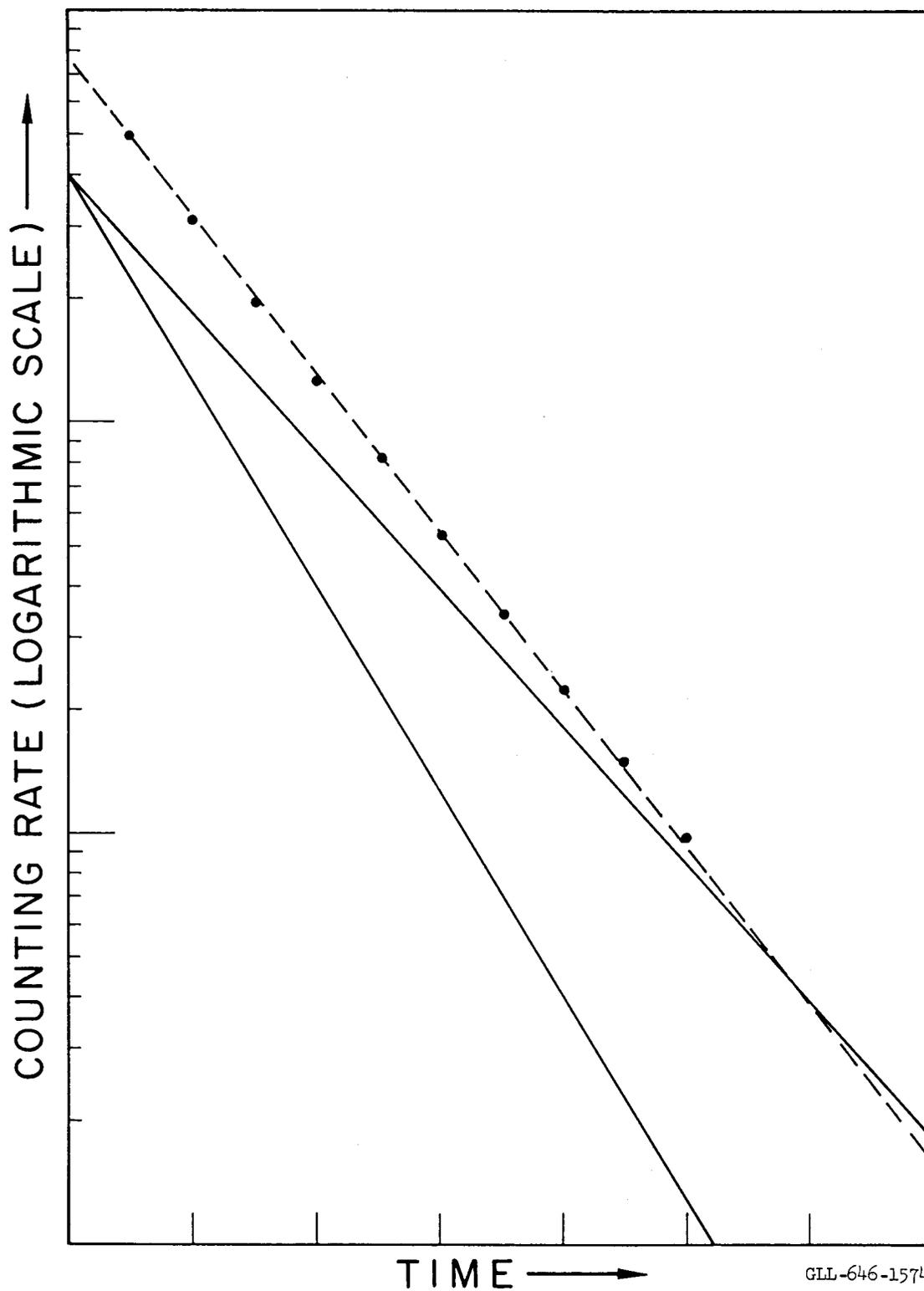
The simplest approach is to plot the data on semilogarithmic paper, $\log(n)$ vs t as before; multicomponent systems will usually give curved lines if the data are well enough distributed in time. If, however, the data do not cover at least two or three half-lives of the apparent principal component, an erroneous conclusion will often be drawn. (See Fig. VI.1.)

A more sensitive test may be applied by performing a least-squares analysis on the data, using the apparent half-life; if two components are actually present, a systematic pattern of deviations will appear, and the χ^2 test will indicate a poor fit.

¹³ J. Tobailem, *J. Phys. Radium* 16, 48 (1955).

¹⁴ R. N. Wonless and H. G. Thode, *Can. J. Phys.* 31, 517 (1953).

¹⁵ C. G. Campbell, private communication, cited in F. Brown and G. R. Hall, *Nuovo Cimento Suppl. (Ser. 10)* 6, 283 (1957).



GLL-646-1574

Fig. VI.1. Decay curve of a two-component system, with equal amounts of two species whose half-lives differ by a factor of 1.5. Deviation from a single-component decay curve of intermediate half-life (dashed curve) is barely perceptible. This illustrates the difficulty of resolving two species with similar half-lives.

A still more sensitive method for detecting the presence of more than one species is based on differential counting. If the gamma spectrum of the sample is followed as a function of time, its shape may often be seen to change with time; a simpler way of accomplishing the same objective is to follow the decay of the sample on both beta and gamma detectors or through various thicknesses of absorbing material. If the sample contains more than one component, the counting rates under different conditions will not all show the same time-dependence; indeed, if the counting conditions can be made sufficiently selective, it is sometimes possible to observe the decay of one or more species essentially pure. In this fortunate case, a good estimate of the half-lives of such species can be obtained. In any case, if the presence of more than one component is suspected, counting should be continued until the activity of the sample has decayed to a small fraction of the original activity if possible.

When there are two or more long-lived components it is not practical to follow the decay of the sample through many half-lives; if the components are not isotopic it is frequently possible to distinguish between them by careful chemical separations. If they are isotopic, the use of a mass-separator should be considered.

5. Known Multicomponent Systems - Evaluation of λ

If we are dealing with a system which is known to contain certain components, and only those components, the evaluation of the λ is possible by several different methods. One usually has approximate values for the λ , which may be obtained graphically as described in Chapter II, Section 2; these approximate values may then be improved by calculation. A direct solution in the manner of Chapter III is not possible, since the forms of the equations are not linear if the λ are treated as unknowns; if, however, Λ_i (the true λ_i) are considered to differ from the approximate λ_i by a small amount ϵ_i so that

$$\Lambda_i = \lambda_i + \epsilon_i \tag{VI.9}$$

then we can say that

$$e^{-\Lambda_i t_j} = e^{-(\lambda_i t_j + \epsilon_i t_j)} = e^{-\lambda_i t_j} \cdot e^{-\epsilon_i t_j}, \tag{VI.10}$$

$$e^{-\Lambda_i t_j} \approx (1 - \epsilon_i t_j) e^{-\lambda_i t_j}, \quad (\text{VI.11})$$

and in each datum equation,

$$x_i e^{-\Lambda_i t_j} = x_i e^{-\lambda_i t_j} - (x_i \epsilon_i) \cdot t_j e^{-\lambda_i t_j}. \quad (\text{VI.12})$$

We therefore replace each unknown x_i with a pair of unknowns, x_i and $\epsilon_i x_i$, and solve as for a system with twice the number of components.* The ϵ_i are evaluated by division of the $\epsilon_i x_i$ by the x_i ; the original values of the λ_i are corrected by adding the ϵ_i , and the process repeated. If the initial values of the λ_i are sufficiently close to the true values so that equation VI.11 is a good approximation, the iterations will usually converge and the calculated values of the ϵ_i will approach zero.

This technique is laborious at best, and beset with difficulties; if the original values of the λ are not sufficiently good, the iteration will diverge instead of converging. It is also possible to "overshoot" the true value of the λ and encounter a situation where the successive values of a given ϵ_i oscillate about zero without decreasing.

Another approach is to consider the value of χ^2 and its dependence on the λ , as derived in Chapter V, equation V.23. If χ^2 is greater than its expected value, one may make the approximation that the deviation of χ^2 from its expected value is

$$\Delta(\chi^2) \approx \sum_{i=1}^n \frac{\partial(\chi^2)}{\partial \lambda_i} \cdot \Delta \lambda_i \quad (\text{VI.13})$$

and select a set of $\Delta \lambda_i$ which will reduce χ^2 to a minimum value. Any systematic scheme for selecting such a set may be used.

A simple scheme is to assume that the deviation of χ^2 from its expected value is due to equal contributions from each λ , so that we select

$$\Delta \lambda_i = \frac{(\chi^2)}{n \left[\frac{\partial(\chi^2)}{\partial \lambda_i} \right]} \quad (\text{VI.14})$$

* The coefficients of the $(\epsilon_i x_i)$ unknowns in the least-squares solution method are of course $t_j e^{-\lambda_i t_j}$ rather than simple exponential expressions.

Another simple scheme is to assume that each λ is in error by a fixed percentage, so that

$$\Delta\lambda_i = \frac{\lambda_i \Delta(\chi^2)}{\sum_{i=1}^n \lambda_i \frac{\partial(\chi^2)}{\partial\lambda_i}} . \quad (\text{VI.15})$$

Still another scheme is to weight the change in each λ_i by the magnitude of the sensitivity of χ^2 to that λ_i , i. e., to change those λ_i to which χ^2 is most sensitive; this gives

$$\Delta\lambda_i = \frac{\frac{\partial(\chi^2)}{\partial\lambda_i} \cdot \Delta(\chi^2)}{\sum_{i=1}^n \frac{\partial(\chi^2)}{\partial\lambda_i}} . \quad (\text{VI.16})$$

Of these schemes, VI.16 is probably better than VI.15 which is better than VI.14.

If any of the λ 's of the system are known precisely, these should not be varied in the iteration scheme. This situation may occur when it is desired to identify an impurity which is observed in a sample containing known components.

6. Unknown Multicomponent Systems

If nothing more is known about the sample than the raw data, consisting of a set of counting rates (and their variances) measured at various known times, the problem of resolution is in the general case extremely difficult. Two general methods of approach have been proposed.

The first method, which has occasionally been used,¹⁶ is a "brute force" method in which one assumes in turn n , $n + 1$, $n + 2$, . . . components, evaluates their half-lives by an iterative procedure based on initially assumed values as outlined in the previous section, and then tests, each time, the statistical improvement in the fit to the data resulting from the addition of the latest component. When the improvement upon adding an additional component is no longer significant, one concludes that the latest added component is spurious and that the system was completely and correctly resolved in the

¹⁶Keepin, Wimett, and Ziegler, J. Nucl. Energy 6, 1 (1957).

previous attempt. It goes without saying that such a method is extremely laborious; it is also somewhat liable to mistake two components with similar half-lives for a single component. The mathematical details of the method are discussed by Householder.¹⁷

A more elaborate method has been described by Gardner et al.¹⁸ Their method is based on integral transforms. In brief, they attempt to determine a function $g(\lambda)$ which is a sum of delta functions, each multiplied by an appropriate amplitude; then if the activity of the sample is represented by $f(t)$,

$$f(t) = \int_0^{\infty} g(\lambda)e^{-\lambda t} d\lambda. \quad (\text{VI. 17})$$

By a suitable transformation of variables, this equation is converted to the complex plane; $f(t)$ then assumes the form of a Fourier transform of $g(\lambda)$. By appropriate integrations, the inverse transformation can be performed and a plot of $g(\lambda)$ vs λ is produced. (See Fig. VI.2.)

This method has the great advantage that the number of components appears as one result of the analysis. The actual plot is expected to resemble a spectrum, with a peak at each value of λ corresponding to a component. The area of the peak is related to the abundance of the component. In actual practice, the situation is more complicated; one usually does not possess data which can be integrated to infinite time, and integration of numerical data is subject to errors. Both these effects contribute to the appearance of numerous small peaks, or "error ripples," which greatly confuse the spectrum. Many problems remain to be solved in the actual application of the method.

VII. DEAD TIME AND COINCIDENCE CONSIDERATIONS

1. Nature of the Problems

Certain statistical problems arise from the physical limitations of the counting equipment itself. In general a single radiation event is detected by some physical system which is in a metastable state - a phototube with a potential across the electrodes, for example, or a wire at high potential in a gas. The event "triggers" the transition of this metastable state to a state

¹⁷ A. S. Householder, ORNL-455 (Unclassified).

¹⁸ D. G. Gardner, J. C. Gardner, G. Laush, and W. W. Meinke, J. Chem. Phys. 31, 978 (1959).

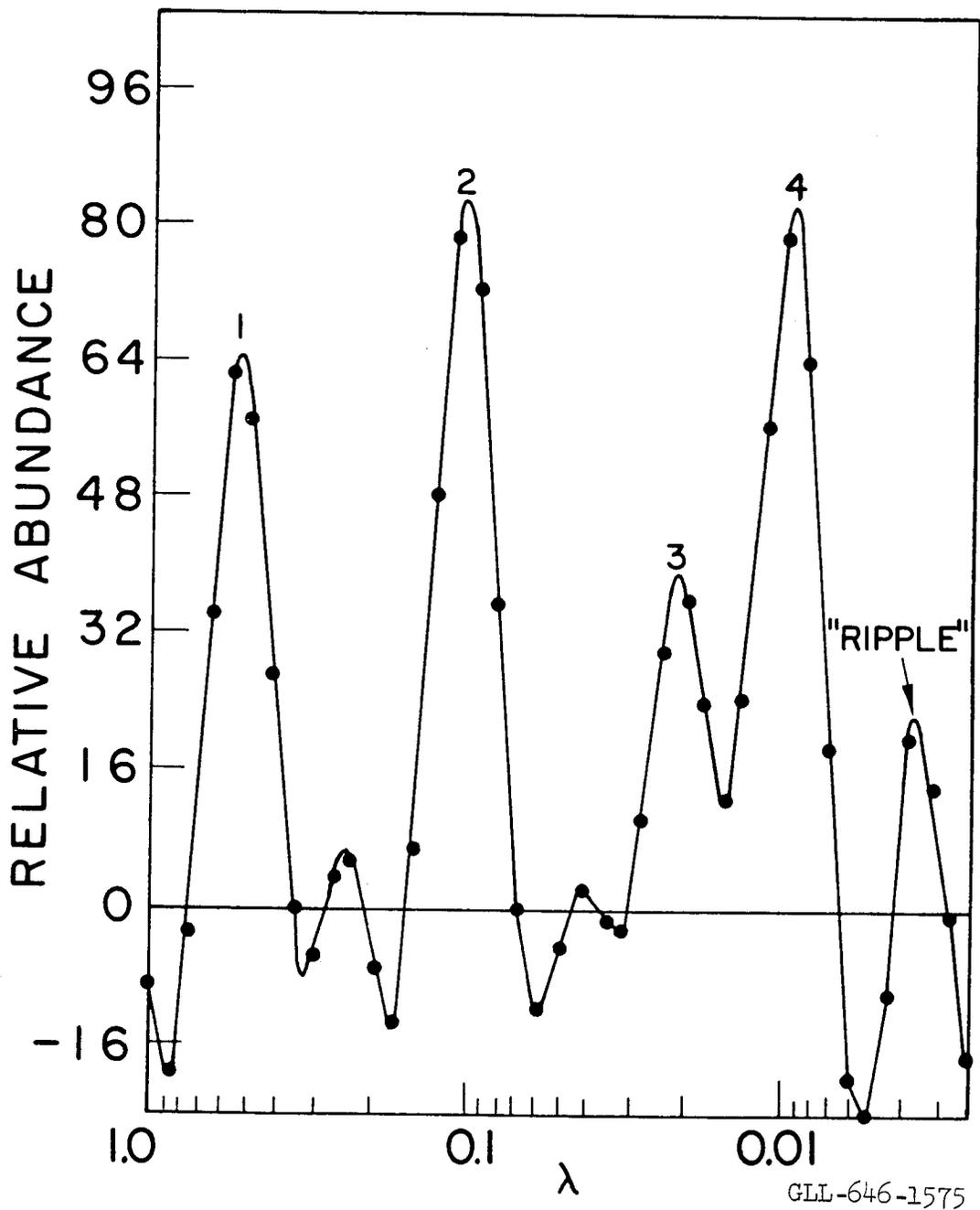


Fig. VI.2. Result of the method of Gardner et al. applied to a set of synthetic data having four components. The four prominent peaks (numbered 1-4) represent components; minor error ripples about the zero-line are inherent in the method. The large ripple at the right-hand end represents failure to obtain data out to infinite time. (Reproduced from Gardner et al., J. Chem. Phys. 31, 978 (1959), Fig. 6, p. 983.)

of lower potential energy, and the energy liberated in the transition is recorded. The system is then returned to its metastable state to await the next event. This process takes time; while the system is reacting to an event it is not capable of reacting to a second event. Two events occurring within a certain time, then, will be recorded as one event; if the system responds in a manner proportional to the energy of the arriving radiations, two events occurring within the critical time will record as a single event having the total energy of the two events. For counting circuits, this critical time is usually spoken of as the "dead time"; for coincidence circuitry, as the "resolving time."

In this chapter we shall examine the effect of the dead time on the observed counting rate; the effect of scaling circuits on interval between pulses into a mechanical register; the measuring of true coincidences vs fortuitous coincidences; determination of counting efficiency by coincidence methods; and the effect of coincident events on energy-discriminated counting circuits.

2. Effect of Dead Time on Observed Counting Rate

Rainwater and Wu⁸ have discussed two limiting situations. In one, a pulse in the counter causes it to be insensitive for a time τ ; a second event arriving during this time does not record and does not extend the dead time. In the second case, a second event arriving during the dead time does extend the dead time to a time τ after the arrival of the second (undetected) event.

Let us consider the first situation. Let the "true" counting rate be N per unit time and the measured counting rate be n per unit time. Then since each recorded count implies a dead time of τ , which is not affected by the unrecorded events, the total time lost per unit time is $n\tau$. The number of events occurring during this lost time is $Nn\tau$, so that

$$N - n = Nn\tau. \tag{VII.1}$$

Consequently, for known dead time,

$$N = \frac{n}{1 - n\tau} \tag{VII.2}$$

is fitted to

$$\sum_y x_y e^{-\lambda_y t_i}.$$

Since

$$n = \frac{N}{1 + N\tau}, \quad (\text{VII.3})$$

it can be seen that the maximum counting rate in this type of instrument, as N increases without limit, is $1/\tau$.

In the second case, where the dead time is extended by τ for each event, whether recorded or not, the analysis is slightly more elaborate. One can see that a second count will be recorded only if it follows the first count by a time interval greater than τ . We must consider the probability that the interval between events, arriving at a rate N per unit time, is at least as great as τ .

The expected number of events during an interval τ is $N\tau$. If Poisson statistics applies, the probability that the actual number is zero, that is, that a second event does not arrive within a time τ , is given by

$$P_0(n) = \frac{(N\tau)^0 e^{-N\tau}}{0!} = e^{-N\tau}. \quad (\text{VII.4})$$

Consequently, the fraction of the time that the counter is "live" is $e^{-N\tau}$; the observed counting rate will be

$$n = N e^{-N\tau}, \quad (\text{VII.5})$$

which for small values of $N\tau$ is approximately equivalent to equation VII.3. For large values of $N\tau$, n in equation VII.5 goes to zero - the counter is "blocked" or saturated. The maximum counting rate can be seen to occur at $N\tau = 1$, where n is equal to $1/\tau e$.

In actual practice, the observed behavior of a system is usually intermediate between the two extremes. Unless a great deal of effort is expended on checking the behavior of the system at high count rates, it is unwise to operate in a region where $N\tau$ is appreciable compared to 1, even if the "dead time" is reputedly known, since when equation VII.5 diverges appreciably from equation VII.3, it is not certain what the actual correction should be unless the actual behavior of the system has been carefully examined.

3. Measurement of Dead Time

There are several general methods of measuring counter dead time. The simplest method (in principle) involves the preparation of a series of long-lived standard samples whose counting rates relative to each other are known a priori; the observed ratios between counting rates are then used with equation VII.2 to give the dead time. Since

$$N_1 = \frac{n_1}{1 - n_1 \tau}, \quad (\text{VII.6a})$$

$$N_2 = \frac{n_2}{1 - n_2 \tau}, \quad (\text{VII.6b})$$

then combining these two equations and solving for τ ,

$$\tau = \frac{n_1/n_2 - N_1/N_2}{n_1(1 - N_1/N_2)} = \frac{n_2/n_1 - N_2/N_1}{n_2(1 - N_2/N_1)}. \quad (\text{VII.7})$$

A set of suitable samples may be prepared by pipetting various carefully measured aliquots of a suitable tracer onto identical backings, making sure that the size and location (on the backing) of each spot is identical with all the others, and that the amount of solid material present is constant. If the samples are not identical in all respects, errors in positioning within the counter and nonidentity of sample scattering and absorption on the various samples will perturb the result. It should be noted that, since evaluation of dead time by a relationship like equation VII.2 is only valid for $N\tau \ll 1$, we are attempting to measure a small deviation from ideal behavior; such a measurement is equivalent to measuring a small difference between large numbers, which (refer to Theorem IV, Chapter III) is subject to large variance if the numbers are statistically independent, as they are in this case.

Another common technique based on the same approach involves the preparation of a divided sample, so mounted that either portion may be inserted in the counter, or both may be inserted together without either portion interfering with the other. If n_1 is the observed counting rate from one portion, n_2 is the observed counting rate for the other, and n_{12} is the observed counting rate for both samples together, we have for the true counting rates

$$N_1 = \frac{n_1}{1 - n_1\tau}, \quad (\text{VII.8a})$$

$$N_2 = \frac{n_2}{1 - n_2\tau}, \quad (\text{VII.8b})$$

$$N_{12} = \frac{n_{12}}{1 - n_{12}\tau} = N_1 + N_2. \quad (\text{VII.8c})$$

Substituting VII.8a and b into VII.8c, we get

$$\frac{n_{12}}{1 - n_{12}\tau} = \frac{n_1}{1 - n_1\tau} + \frac{n_2}{1 - n_2\tau} \quad (\text{VII.9})$$

which upon simplifying and solving gives

$$\tau = \frac{n_1 n_2 \pm \sqrt{n_1 n_2 (n_{12} - n_1)(n_{12} - n_2)}}{n_1 n_2 n_{12}}. \quad (\text{VII.10})$$

It can be shown that choice of the upper sign will result in a negative answer for N_1 or N_2 ; consequently we have

$$\tau = \frac{n_1 n_2 - \sqrt{n_1 n_2 (n_{12} - n_1)(n_{12} - n_2)}}{n_1 n_2 n_{12}}. \quad (\text{VII.11})$$

This technique appears simple in theory; in practice, extreme care must be used to ensure that each portion of the split sample is counted separately in exactly the same position it occupies when the portions are counted together. The term under the radical sign does not differ greatly from $n_1 n_2$; we are again faced with a small difference between large numbers. Repeated measurements should be made with different sets of split samples; counts should be taken with high precision, recording as many events as possible during each of the three measurements.

A second method for measuring dead time involves the preparation of a strong sample of a short-lived single species whose half-life is very well known. (C^{11} , with a half-life of 20.5 minutes, is a convenient species for the purpose.) If the initial activity of the sample is high enough so that dead time losses are appreciable, then the activity will not decay exponentially with time; if n is the observed counting rate, and the true counting rate decays as a single-component exponential, then

$$n_i = \frac{N_0 e^{-\lambda t_i}}{1 + N_0 \tau e^{-\lambda t_i}} \quad (\text{VII.12})$$

and if enough points are available the parameters N and τ can be evaluated either by a least-squares method or graphically.

A simple graphical approach to give an approximate answer may be derived by reciprocating equation VII.12 to give

$$\frac{1}{n_i} = \frac{1}{N_0} e^{+\lambda t_i} + \tau. \quad (\text{VII.13})$$

If $1/n_i$ is plotted against $\exp(+\lambda t_i)$ a straight line should be obtained with a slope of $1/N_0$ and an intercept of τ . The same form of the equation can readily be used for a least-squares solution, providing that the variances of the $1/n_i$ can be derived. This is done as follows: rewriting VII.13 as

$$\frac{1}{n_i} = \frac{1}{N_i} + \tau, \quad (\text{VII.14})$$

then

$$\text{var}\left(\frac{1}{n_i}\right) = \text{var}\left(\frac{1}{N_i}\right) = \frac{1}{N_i^2} \text{var}(N_i) = \frac{N_i}{\Delta t_i \cdot N_i^2}, \quad (\text{VII.15})$$

$$\text{var}\left(\frac{1}{n_i}\right) = \frac{1}{N_i \cdot \Delta t_i}, \quad (\text{VII.16})$$

where Δt_i is the counting interval (see Chapter V, equation V.7). Unfortunately, N_i involves the parameter τ ; we are again involved in an iterative calculation, where a first approximation for τ is obtained in some reasonable fashion; the weighting factors for the least-squares solution are computed, using this initial value; a new value for τ is obtained by least-squares analysis; if this value is appreciably different from the initial value, the new value of τ is used to recompute improved weighting factors and the procedure is repeated until the value of τ converges to its best value. Since $1/n_i$ is expected to be considerably greater than τ if the simple concept of dead time is to have meaning, the convergence should be quite rapid.

One may also estimate dead time by electronic methods; if the output of the counter is used to trigger the timed sweep of an oscilloscope, and the

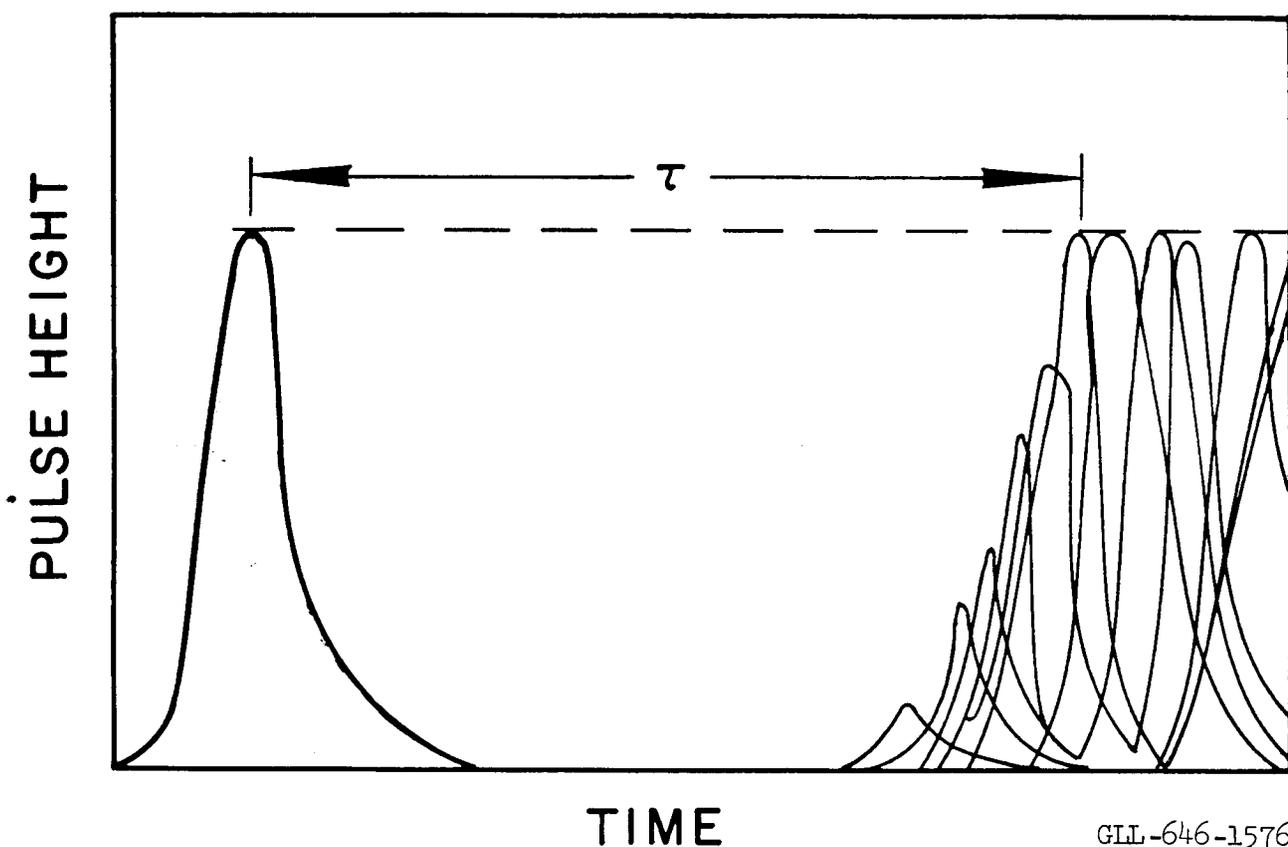
counter pulse is displayed vertically on the scope, then if a highly active source is put into the counter the scope screen will display a pattern like that in Fig. VII.1 (in the ideal case). In Fig. VII.1 the first pulse is displayed at the left end of the scope trace; pulses which follow during the sweep of the scope will appear at random at any point of the sweep after termination of the dead time. If a large number of sweep patterns are superposed, as for example by taking a time-exposure photograph of a number of sweeps, the dead time may be measured by the width of the gap between the first pulse and the earliest second pulse. τ is sometimes a function of pulse height.

Electronic circuits are becoming available which measure counting time intervals in terms of "live time." When such equipment is available, or where the dead time is so short or the counting rate so low that dead time losses are negligible, much trouble may be avoided.

4. Scaling Circuits - Avoiding Losses in Slow Components

If the counting instrument embodies a mechanical register or other slow-acting component (very few mechanical registers will accept pulses separated by only 0.1 second) serious trouble may arise. Even when the average interval between pulses is long enough to permit response of the register, the statistical distribution in time of the pulses implies that a certain fraction of the pulses will occur too close together to be resolved. It is customary to use an electronic scaling circuit between the counter and the register, so that for every m events in the counter only one pulse is fed to the register. This obviously reduces the register counting rate, and increases the average interval between pulses, by a factor of m ; we shall show, following the treatment of Rainwater and Wu,⁸ that there is an additional beneficial effect gained from the regularizing action of such a scaling circuit on the output pulse rate. In other words, the distribution of intervals between output pulses from a scale-of- m circuit is narrower than that from a counter counting at the same rate, i. e., with the same average spacing between pulses. This implies that a given register might be satisfactory, used with a counting rate of mN and a scale-of- m circuit, even though the losses due to slow register response would be unacceptable if the same register were used with a counting rate of N and no scaling circuit.

We wish to consider the probability $q_m(x) dx$ that the actual interval between two output pulses from the scaling circuit lies between x and $x + dx$,



GLL-646-1576

Fig. VII. 1. Schematic representation of the measurement of counter dead time τ using an oscilloscope. Pulses from the counter are used to trigger the horizontal sweep of the oscilloscope. Events which occur within the sweep time appear as peaks in the right-hand portion of the trace. Repeated traces are recorded, e. g. by a long-exposure photograph; counter dead time results in a blank region in which no record pulses appear.

where x is time measured in units of the average spacing between output pulses, i. e., if the output pulse rate is N , and the input pulse rate is mN , then for a time t , x is equal to Nt . We shall start by considering the probability $Q_m(x)$ that an output pulse will occur within a time t of the preceding output pulse.

During the time interval t the scaling circuit will receive some number n of input counts. Let us assume that Poisson statistics applies; then the probability of n input counts is given by

$$P_n = \frac{\overline{E(n)}^n e^{-E(n)}}{n!} \quad (\text{VII.17})$$

and the expected number of counts is

$$E(n) = mNt = mx. \quad (\text{VII.18})$$

An output pulse will have been obtained if $n \geq m$. Consequently the probability of obtaining an output pulse is the sum of all P_n for $n \geq m$; or

$$Q_m(x) = \sum_{n=m}^{\infty} P_n(mx) = 1 - \sum_{n=0}^{m-1} P_n(mx). \quad (\text{VII.19})$$

The probability that an output pulse will occur during the interval between x and $x + dx$ is the change in $Q_m(x)$ when x increases from x to $x + dx$; in other words

$$q_m(x) = \frac{d}{dx} [Q_m(x)]. \quad (\text{VII.20})$$

Substituting equations VII.17 and VII.18 in equation VII.19, we get

$$q_m(x) = \frac{d}{dx} \left[1 - \sum_{n=0}^{m-1} \frac{(mx)^n e^{-mx}}{n!} \right], \quad (\text{VII.21})$$

$$q_m(x) = m \cdot \frac{(mx)^{m-1} e^{-mx}}{(m-1)!} \quad (\text{VII.22})$$

If this is written

$$q_m(x) dx = \frac{(mx)^{m-1} e^{-mx}}{(m-1)!} d(mx) \quad (\text{VII.23})$$

it is apparent that it corresponds to the form of a Poisson distribution, with an expected value of mx equal to $m - 1 + 1 = m$, so that $E(x) = 1$ as might be expected; the point of principal interest is that the width of this distribution, as measured by its variance, is also m ; so that

$$\text{var}(mx) = m^2 \text{var}(x) = m, \quad (\text{VII.24})$$

$$\text{var}(x) = \frac{1}{m}, \quad (\text{VII.25})$$

and the width of the distribution in x varies inversely as m . Figure VII.2 shows the shapes of the distributions for several values of m , including $m = 1$ (no scaling). The curves are normalized so that the area under the curve for $x \leq x_1$ is the probability that the interval between consecutive output pulses is $x \leq x_1$. The very small areas under the curves for small values of x when a scaling circuit is used should be compared to the large area under the curve for $m = 1$, i. e., no scaling; this clearly demonstrates the advantage to be gained by the use of even small scaling factors.

5. Measurement of Coincident Events

It is sometimes desired to measure events which are "coincident" in the sense that they occur, one in each of two counters, within a certain time of each other. It is of course meaningless to require that the two events occur at precisely the same instant; the uncertainty principle alone implies a certain imprecision in the time of occurrence of events of finite energy. In addition, any real physical instrument for measuring coincident events has a finite resolving time τ ; events occurring within a time τ of each other will be recorded as coincident.

Events measured as coincident may be divided into two classes: "truly coincident" events which are related in some physical way, such as two gamma photons emitted in cascade without noticeable delay in the intermediate state; and "chance coincident" events, which arrive together at the detection system within the time τ through pure chance. Fortunately, it is possible to calculate the rate at which chance coincident events will occur; the gross rate of

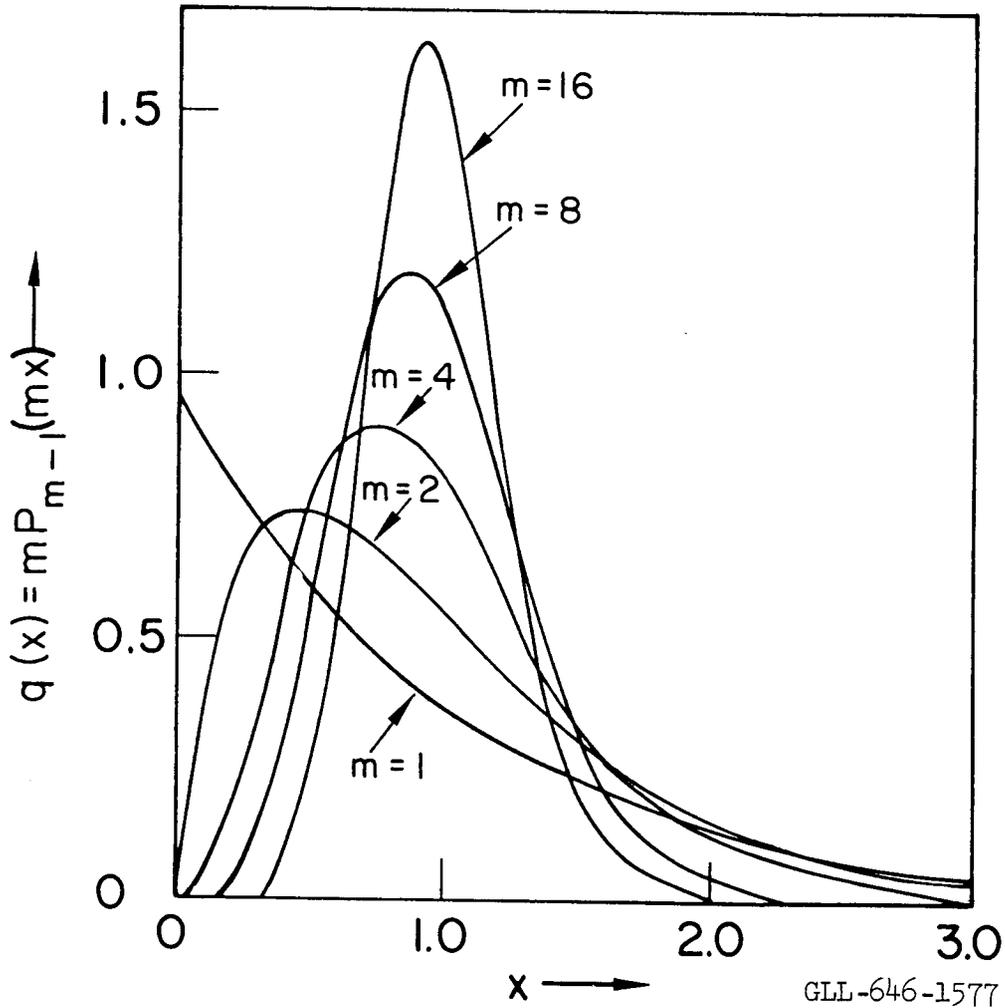


Fig. VII.2. Regularizing action of scaling circuits. $q(x) dx$ represents the probability that the time interval between two pulses to the register will lie between x and $x + dx$, where x is a fraction of the mean interval. The probability that the interval will be less than the dead time, a , of the register is represented by the area under the curve from $x = 0$ to $x = a$; m is the scaling factor in the scaling circuit. It can readily be seen that even rather small scaling factors are very helpful in reducing losses. (Reproduced from Rainwater and Wu, *Nucleonics*, October 1947, p. 60; Fig. 3, p. 69.)

coincident events can then be corrected to give the true coincidence rate. The calculation is made as follows: Let us assume that, as is almost always true, the true coincidence counting rate is very small compared to the counting rate in either counter. If the counting rate in one counter is n_1 , and the counting rate in the second counter is n_2 , then the probability of a count in one counter during the interval τ is proportional to both n_2 and τ ; it is, in fact, $n_2\tau$. The usual circuit for coincidence measurements is so arranged that a coincidence will be recorded if a count in counter 2 follows a count in counter 1 within an interval τ , but not the reverse; in such a case, the chance coincidence rate will be $n_1n_2\tau$. If n_1 and n_2 are recorded and τ is known, the chance coincidence rate can readily be calculated.

In planning an experiment it is desirable to make an estimate of the ratio of true to chance coincidences. For this we need an estimate of the counting efficiency of each counter. If the true disintegration rate is N , and the efficiencies in each counter are η_1 and η_2 where efficiency in this case is taken to include branching ratios, decay schemes, and selective counting response, then the true coincidence rate can be seen to be $\eta_1\eta_2N$ and the chance coincidence rate can be seen to be $\eta_1\eta_2N^2\tau$. The ratio of "true" to "chance" coincidences will then be $1/N\tau$. It is therefore desirable to operate with as short a resolving time as possible, which seems obvious a priori; what is not so obvious is that it is also advantageous from the point of discriminating against false coincidences to operate with a sample containing as little activity as possible.

If this is done, it then becomes highly desirable to increase the solid angle intercepted by the counters to the largest possible value so that the observed gross counting rate will be as high as possible, in order to obtain good statistical reliability in a reasonable experimental period. We are now in some danger of violating our simplifying assumption; if we consider a limiting case in which every count in the first counter gives rise to a true coincident count in the second counter, it is obvious that while τn_1n_2 is not zero the actual false coincidence rate is truly zero.

The evaluation of true versus false coincident counts is somewhat more complicated if we omit the simplifying assumption. Let us assume a "true coincident" counting rate of n_{12} ; then there will be n_{12} counts per unit time in both counters 1 and 2 which should not be included in the evaluation of the

false rate. If g is the gross observed coincident count rate, and f is the false coincident rate, then

$$g = f + n_{12} \quad (\text{VII.26})$$

and

$$f = (n_1 - n_{12})(n_2 - n_{12})\tau \quad (\text{VII.27})$$

so that, combining VII.26 and VII.27,

$$n_{12} = \frac{\tau(n_1 + n_2) - 1 + \sqrt{[1 - \tau(n_1 + n_2)]^2 + 4\tau(g - \tau n_1 n_2)}}{2\tau} \quad (\text{VII.28})$$

This can be transformed to

$$n_{12} = \left[\frac{1 - \tau(n_1 + n_2)}{2\tau} \right] \left[-1 + \sqrt{1 + \frac{4\tau(g - \tau n_1 n_2)}{[1 - \tau(n_1 + n_2)]^2}} \right] \quad (\text{VII.29})$$

which can be approximated by

$$n_{12} \rightarrow \frac{g - \tau n_1 n_2}{1 - \tau(n_1 + n_2)} \text{ as } \left[\frac{4\tau(g - \tau n_1 n_2)}{[1 - \tau(n_1 + n_2)]^2} \right] \text{ becomes } \ll 1. \quad (\text{VII.30})$$

Since for any reasonable system $\tau(n_1 + n_2) \ll 1$, then the simplifying assumption does not give rise to serious error unless $4\tau n_{12}$ is appreciable compared to 1.

6. Variance of Coincident Event Rate

If one considers a system such as the above, but containing a sample of such a nature that there are no true coincidences, then it becomes apparent that the gross coincidence count rate observed must be the actual false rate; that the expected value of the difference between the observed gross rate and the calculated false rate is zero; and that the variance of the difference between the gross and the calculated false rates is just the variance of the false coincidence rate. This variance can be computed from theory.

Since we know that the expected value of the false rate, f , is given (to a good approximation) by

$$f = \tau n_1 n_2, \quad (\text{VII.31})$$

then

$$\text{var}(f) = \tau^2 \text{var}(n_1 n_2) = \tau^2 n_1 n_2 \left(\frac{n_1 + n_2}{t} \right) \quad (\text{VII.32})$$

where t is the duration of the count (see Chapter V, equation V.7). Since for true coincidences Poisson statistics should apply,

$$\text{var}(n_{12}) = \frac{n_{12}}{t} \quad \text{if } f = 0; \quad (\text{VII.33})$$

then from the preceding discussion

$$\text{var}(n_{12}) = \frac{g - f}{t} + \tau^2 n_1 n_2 \left(\frac{n_1 + n_2}{t} \right). \quad (\text{VII.34})$$

This variance is then used for least-squares weighting.

7. The Effect of Coincident Events on Discriminated Counts

Several varieties of counters may be made to give an output pulse with amplitude nearly proportioned to the energy of the event giving rise to the pulse. Such counters, in conjunction with suitable discriminating or analyzing circuits, may be used to give particle energy spectra, or to count radiations of a particular energy in the presence of a variety of others. In this situation coincident events may occur in which two particles enter the counter within the resolving time of the system; the amplitude of the resulting pulse is then proportional to the sum of the energies of the two particles.

When one is examining a sample which gives rise to a pair of truly coincident events - for example, two gamma photons in cascade - then if the probability of one event recording in the counter is p_1 , and the probability of the other recording in the counter is p_2 , the probability of both events recording in the counter is $p_1 p_2$. Consequently, the ratio of coincident events, giving rise to a "sum" peak with energy $E_1 + E_2$, to single events of one type (e. g., type 1) will be

$$\text{Sums/single} = \frac{p_1 p_2}{p_1 - p_1 p_2} = \frac{p_2}{1 - p_2} \approx p_2. \quad (\text{VII.35})$$

p_1 and p_2 are simply the counting efficiencies of the system for the types of radiation in question; consequently, for a given decay rate, p_2 is proportional to the singles counting rate of the second type of event. Typical counting efficiencies are $\geq 10^{-3}$ in most cases and may even approach unity in special arrangements.

Let us now consider a sample which does not give rise to truly coincident events, in a system with resolving time τ . The relationship between count rates for single events and count rate for false coincidences is the same as that derived in the previous section; the ratio between "false sums" and singles of type 1 will be

$$\text{Sums/singles} = \frac{n_1 n_2 \tau}{n_1} = n_2 \tau = p_2 N \tau \quad (\text{VII.36})$$

(where N is the disintegration rate), which is again proportional to the count rate of the other type of event. In this case the two types of events may be photons of different energy originating from different atoms, giving rise to a "sum peak" of energy $E_1 + E_2$, or photons of the same energy, giving rise to a "sum peak" of energy $2E_1$, in which case N_2 and N_1 are equal. It will be seen that there are differences between true and false sum peaks. The size of the proportionality factor is larger in general for true coincidences; but much more important is the fact that p_2 depends only on solid angle, counter characteristics, and to a slight extent on sample mounting, so that the ratio of sums to singles in a particular counting arrangement depends only on counting position, i. e., solid angle, for true coincidences, but depends on absolute disintegration rate for false coincidences. The sums-to-singles ratio will therefore change with time as the sample decays and the absolute disintegration rate changes.

A point that should be made is that a nuclear species which has two gamma photons in coincidence may actually have a single gamma photon whose energy is the sum of the two coincident gammas; the decay scheme for such a species is drawn in Fig. VII.3. The gamma spectrum of such a species will contain a peak of energy $E_1 + E_2$ which will arise from three sources: single events corresponding to the "crossover" transition from the upper level to the ground state; true coincident counts of the two gammas in cascade; and false coincidences resulting when one of the cascade gammas strikes the counter in fortuitous coincidence with the cascade gamma of complementary energy from another nucleus. In attempting to demonstrate the real existence of the "crossover" transition, and to measure its abundance, due allowance must be made for contributions to the gamma peak from the other sources.

It should also be noted that a multicomponent sample is sometimes followed for decay in a proportional counter with a discriminator set to block

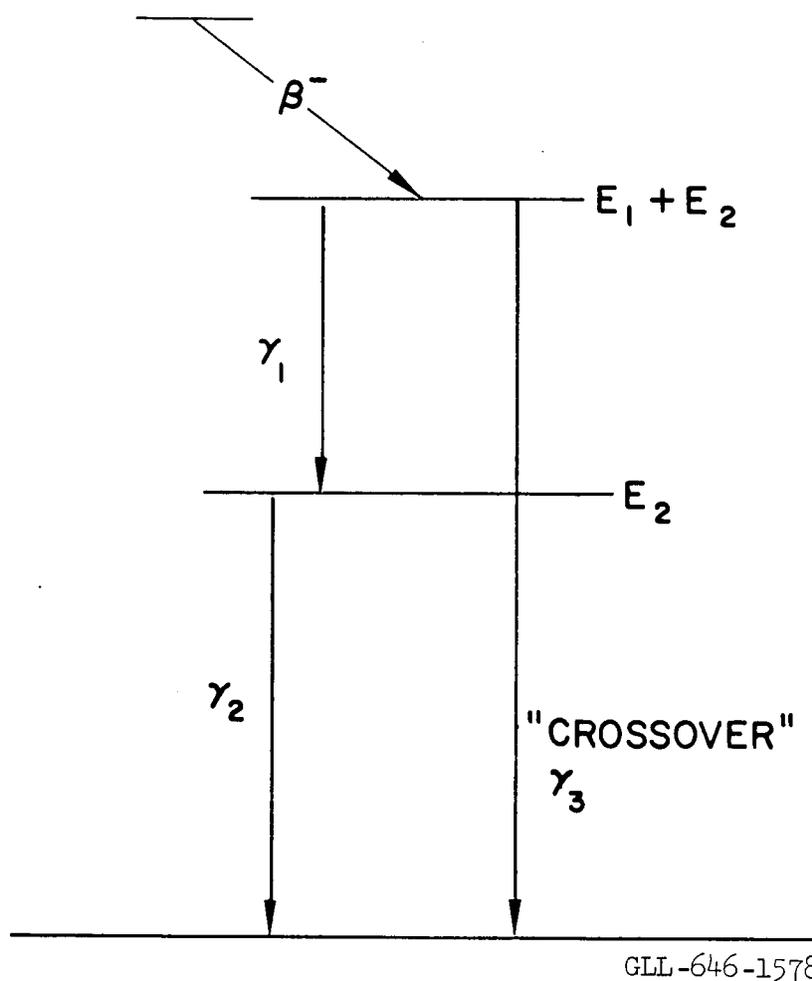


Fig. VII.3. Decay scheme illustrating a possible difficulty in interpreting "sum peaks" in gamma spectra. A peak will appear in the spectrum of such a nuclide at the gamma energy $E_1 + E_2$. This arises from three sources: a "true coincidence" when the photons γ_1 and γ_2 both strike the crystal; a "crossover" when the nuclide decays directly from the upper excited state to the ground state; and a "false coincidence" when a photon of energy E_1 from one nucleus and a photon of energy E_2 from another both strike the crystal fortuitously within the resolving time of the instrument. The intensity of the "sum peak" will vary with solid angle and with absolute source intensity in different ways for the three different phenomena.

off a component which gives rise only to relatively low-energy radiations, while recording the higher energy radiations from another component. If there are true coincidences in the undesired species, the "sum peak" may lie above the discriminator level and the unwanted component will appear in the decay curve. If there are no coincident events, but the unwanted component is sufficiently active to give an appreciable number of false coincidences, it will easily be seen that the false coincident rate will depend on the square of the count rate; the decay rate for the false sum peak will be given by

$$R \propto (R_0 e^{-\lambda t})^2 = R_0^2 e^{-2\lambda t} \quad (\text{VII.37})$$

and a component will appear in the decay having precisely half the half-life of the supposedly eliminated species. This effect can be quite startling if not anticipated.

8. Measurement of Absolute Counting Efficiency by Coincidence Techniques

If one has a radioactive species which gives rise to truly coincident events, measured in an experimental arrangement of two counters, it is possible to use the coincidence counting rate to measure the true disintegration rate of the sample. If the counting efficiencies in the two counters, including solid angle and branching ratio if necessary, are η_1 and η_2 , then the counting rates are given by

$$\begin{aligned} n_1 &= \eta_1 N, \\ n_2 &= \eta_2 N, \\ n_{12} &= \eta_1 \eta_2 N + \tau n_1 n_2 \quad (\text{gross coincidence rate}), \end{aligned} \quad (\text{VII.38})$$

and consequently

$$N = \frac{n_1 n_2}{n_{12} - \tau n_1 n_2} . \quad (\text{VII.39})$$

It is of some interest to consider the precision of such a measurement. We have seen (equation VII.34) that

$$\text{var}(n_{12} - \tau n_1 n_2) = \frac{n_{12} - \tau n_1 n_2}{t} + \frac{\tau^2 n_1 n_2 (n_1 + n_2)}{t} \quad (\text{VII.40})$$

and consequently, using Theorems V and VI of Chapter III, we get, where t is the duration of the measurement,

$$\text{var}(N) = \frac{1}{t} \left(\frac{n_1 n_2}{n_{12} - \tau n_1 n_2} \right)^2 \left[\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_{12} - \tau n_1 n_2} + \frac{\tau^2 n_1 n_2 (n_1 + n_2)}{(n_{12} - \tau n_1 n_2)^2} \right]. \quad (\text{VII.41})$$

VIII. GLOSSARY OF SYMBOLS

$ A $	determinant of the matrix A
A^{-1}	inverse of the matrix A
A^t	transpose of the matrix A
A_{ij}	element of the matrix A
a_{ij}	coefficient in a system of linear equations
b, B	total number of background events
B, C, H	matrices
cov(x, y)	covariance of two variates x and y
D	diagonal matrix
$E(x)$	expected value of a variate x
$E(f(x))$	expected value of a function of a variate x
f	false coincidence rate
F_e	extrapolation factor
g	growth factor; gross counting rate, events/unit time; right member vector of the normal equations; gross number of events; gross observed coincidence rate
$g(\lambda)$	function used to resolve unknown mixtures (see p. 107)
• G	vector of right member of normal equations; gross number of events
i	vector with unit components
I	unit matrix
k	number of replicate samples or measurements
K_x	covariance matrix of a variate vector
m	number of measured data points
n	number of atoms in sample; number of components in system; number of observed events; counting rate modified by dead time
N	number of atoms in sample; true counting rate
$n\phi$	integrated flux over the sample
O	zero matrix
p	observed-calculated number; probability that an atom will decay and give rise to a count
$P_x, P_x(u)$	distribution function of x (along coordinate u)
$q_m(x)$	probability distribution function for intervals between events
$Q_m(x)$	probability distribution function for output pulses (integral)
Q_n	special polynomial defined in eq. IV. 25, p. 69

R	vector of residuals
R_g	renormalizing factor (see eq. IV.39, p. 73)
r_j	residual, difference between observed point and calculated curve
S	weighted sum of squares; scaling factor
t	time
t_B, t_G	duration of measurement of background or gross count
$T_{1/2}$	half-life
u	expected value of count
U	upper triangular matrix
V, W	variate vectors
var(x)	variance of x
w	waiting period
w_i, W_i	weighting factors
x	number of atoms <u>or</u> counts per minute present at zero time <u>or</u> a specified time
\bar{x}	average value of x
X	vector with components x_i ; convenient symbol for $E(x)$
y_j	counting rate; weighted counting rate
β	observed background count; instrumental variance factor for background
γ	instrumental variance factor for gross count
Δ	duration of an irradiation
δ_{ij}	delta function: $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ otherwise
Δt	duration of a counting measurement; time between zero time and first count
ϵ	small error
η	counting efficiency; true values of counting rates (weighted)
λ	decay constant = $(\ln 2)/T_{1/2}$
Λ	true decay constant
ξ	true values of the values to be solved for
$\pi_N(z)$	probability that number of atoms in a sample is z
σ	reaction cross section; standard deviation
σ^2, σ_x^2	variance (of x)
τ	dead time; resolving time
ϕ	flux (particles per unit area per unit time)
χ^2	chi-squared (statistical test)

IX. INDEX

	<u>Page No.</u>
Absolute counting:	
measurement of --- efficiency by coincidence techniques	127
precision of --- by coincidence	127
Absorbing:	
radiation scattering and ---	65
Activation analysis	9
Activities:	
use of daughter --- to measure nuclear reactions . . .	13
Activity:	
growth of ---	13
half-life by direct measurement of specific --- . . .	103
measurement of half-life by specific ---	103
Algebra:	
matrix ---	26
of covariances	36
of expected values	36
of linear equations	26
of matrices	26
of variances	36
Amplitude	124
Analysis:	
activation ---	9
improper use of least-squares ---	94
of replicate samples	35,95
of variance	35,95
statistical imprecision of a single radiochemical --- .	95
Analytical:	
variance due to --- technique	96
variations	95

	<u>Page No.</u>
Arithmetic average	40
Assumptions underlying decay equations	3
Average:	31
arithmetic ---	40
best --- value for a given number of replicates	98
best value of an ---	46
estimated measure of the precision of the --- of a number of measurements	42
having the least variance	45
individual quantities which are to be --- are not uncorrelated	45
variance of an ---	40
variance of a weighted ---	47
variance of best --- value for a given number of replicates	98
weighted ---	44,45
weighting factor to --- the replicate results	96
Averaging:	
and sampling statistics	39
measurements of unlike variances	45
of nonindependent quantities	45
replicate values	98
weight in the --- process	44
Avoidance of computer overflow	94
Avoiding:	94
gross errors	94
losses in slow circuit components	117
Background:	76,84
correction for counter ---	76
correction when Poisson distribution is not applicable	76
nonrandom	76
renormalized probability distribution of --- count	76

Bernoulli distribution	
expected value of	65
expected value of a count from observed count for ---	75
standard deviation of ---	66
variance of expected value of a count for ---	75
Blocking of a counter	113
Branched decay	21
Branching:	
decay	19
ratio	19
growth factor in --- decay	20
Buildup toward equilibrium	8
Cascade	124
Chains:	
decay --- with more than two consecutive decays	21
Chance:	
coincident events	120
ratio of true to --- coincidences	122
Chi-square test	88
Chi-squared:	63
test of goodness of fit	60
Circuit:	
avoiding losses in slow --- components	117
distribution of intervals between output pulses from a scale-of-m ---	117
electronic scaling ---	117
regularizing action of a scaling ---	117,121
scaling ---	85,117
Coefficients	26
Cofactor of a matrix element	29
Coincidence:	110
counting of large solid angles	122
expected value of the false --- rate	123
false ---	125,126

	<u>Page No.</u>
Coincidence:	
measurements of absolute counting efficiency by --- techniques	127
precision of absolute counting by ---	127
ratio of true to chance ---	122
true	126
Coincident:	
chance --- events	120
effects of --- events on discriminated counts	124
events	124
measurement of --- events	120
true --- counts	125
true versus false --- counts	122
truly --- events	120
variance of --- event rate	123
variance of the false --- rate	123
Column vector	27
Combination:	
of errors	37
of matrices	27
variance of any linear --- of variates	39
Complex systems:	
mathematical techniques for deriving formulas for decay of	21-24
Component:	4
precision of the amount of each --- present	87
Components:	
avoiding losses in slow circuit ---	117
correction for presence of daughter in parent-daughter systems where both --- have comparable half-lives	13
F-test for number of ---	60
identification of ---	101
number of ---	101
two or more long-lived ---	107

	<u>Page No.</u>
Computational: least-squares --- method	61
Computer overflow: avoidance of ---	94
Confidence:	
limits	33, 34
n-percent --- interval	33
Consecutive:	
decay chains with more than two --- decays	21
decay for parent and daughter with equal half-life	11
disintegrations	10
Constant:	35
decay ---	3
method of evaluating a decay ---	101
Constants:	
partial decay ---	101
Correction:	
actual --- for dead time	113
background --- when Poisson distribution is not applicable.	76
factor for presence of daughter	12
for counter background	76
for presence of daughter in parent-daughter systems where both components have comparable half-lives	13
for presence of daughter in parent-daughter systems where daughter is considerably longer-lived than parent	13
Correlation	35
Count:	
distribution of expected values for a given observed --- for a large sampling	69
distribution of expected values for a given observed --- for a small sampling	68
expected value of a --- from observed --- for Bernoulli distribution	75
expected value of a --- from observed value for Poisson distribution	69

Count:

expected value of net ---	77
gross ---	84
high --- rates	113
probability of disintegrating and producing a ---	64
renormalized probability distribution of background ---	76
variance of an extrapolated --- rate	79
variance of an observed ---	84
variance of expected value of a --- for Bernoulli distribution	75
variance of expected value of --- for a Poisson distribution	69
variance of expected value of net ---	76

Counter:

blocking of a ---	113
correction for --- background	76
measurement of --- dead time with an oscilloscope	118
measuring --- dead time	114
saturation of a ---	113

Counting:

additive variance due to --- equipment	86
coincidence --- of large solid angles	122
definition of --- data	1
differential ---	107
distribution functions of --- results	64
effect of dead time on observed --- rate	112
efficiency	64
geometry of --- arrangement	65
maximum --- rate	113
measurement of absolute --- efficiency by coincidence techniques	127
net --- rate	84
precision of absolute --- by coincidence	127

Counting:

probability that precisely n counts will be recorded during the --- interval	65
rate	4
single photopeaks	87
statistics of --- data	64
total --- rate of a sample	11
variance of net --- rate	84

Counts:

effect of coincident events on discriminated ---	124
extrapolating --- per unit time	81
probability that precisely n --- will be recorded during the counting interval	65
true coincident ---	125
true versus false coincident ---	122

Covariance:

diagonal elements of --- matrix are variances	87
matrix	39, 46, 52
of x and y	33

Covariances:

algebra of ---	36
--------------------------	----

Criteria for rejection of data	92
------------------------------------------	----

Criterion: additional --- for rejecting data	93
--------------------------------------------------------	----

Cross section: reaction --- measurement	83
---------------------------------------------------	----

Crossover:	126
----------------------	-----

transition	125
----------------------	-----

Cutoff: energy ---	87
------------------------------	----

Daughter:	10
---------------------	----

analytical scheme when both parent and --- are formed directly by nuclear reactions	17
-----------------------------------------------------------------------------------------------	----

analytical scheme when parent is much shorter-lived than ---	16
------------------------------------------------------------------------	----

Daughter:

analytical scheme when both parent and --- are formed directly by nuclear reactions	17
analytical scheme when parent is much shorter-lived than ---	16
analytical schemes when half-life of parent is much longer than that of ---	15
consecutive decay for parent and --- with equal half-life	11
correction factor for presence of ---	12
correction for presence of --- in parent-daughter systems where both components have comparable half-lives .	13
correction for presence of --- in parent-daughter systems where --- is considerably longer-lived than parent .	13
use of --- activities to measure nuclear reactions .	13

Dead time: 110,112,113

actual correction for ---	113
effect of --- on observed counting rate	112
estimation of --- by electronic methods	116
estimation of --- by graphical methods	116
measurement of ---	114
measurement of counter --- with an oscilloscope . . .	118
measuring counter ---	114
measuring --- with a divided sample	114
measuring --- using a short-lived species of well-known half life	115

Decay:

assumptions underlying --- equations	3
branched ---	21
branching ---	19
chains with more than two consecutive ---	21
comparison of the --- of the unknown species with that of a species with precisely known half-life.	103
consecutive --- for parent and daughter with equal half-life	11
constant	3
curve	5,6
curve of a two-component system	106
formation of active species by --- of a radioactive precursor	10

Decay:	<u>Page No.</u>
fundamental law of radioactive ---	3
graphical resolution of --- curve	6
growth factor in branching ---	20
least-squares solution to resolve --- data	83
mathematical techniques for deriving formulas for --- of complex systems	21-24
matrix inversion method for resolution of --- curve .	87
method of evaluating a --- constant	101
observing --- over long periods of time	103
of a parent-daughter pair where parent is the longer- lived species	14
of a parent-daughter pair where parent is the shorter- lived species	14
of a parent-daughter system	12
partial --- constants	19
proper weighting factors for use in a least-squares calculation of --- data	87
proper weighting of --- data for a least-squares solution	83
rate for the false sum peak	127
resolution of --- data	83
resolving the --- data	4
scheme	65
two-component --- curve	6
 Decays:	
self-calibrating method for measuring long-lived --- .	105
Definition of counting data	1
Degrees of freedom	43
Depleted target material	9
Destruction of product	10
Detection of multicomponent systems	105
Detector efficiency	9
Determinant of a square matrix	29
 Deviation:	
standard --- of Bernoulli distribution	66
standard --- of Poisson distribution	67
standard --- of normal distribution	68
standard --- of x	33

	<u>Page No.</u>
Deviations:	
number of standard ---	92
Diagonal:	
elements of covariance matrix are variances	87
inverting a --- matrix	63
matrix	29
Difference variance	37
Differential counting	107
Discriminated:	
effect of coincident events on --- counts	124
Disintegrating:	
probability of --- and producing a count	64
Disintegration:	
measurement of true --- rate	127
Disintegrations:	
consecutive	10
Distribution:	
background correction when Poisson --- is not applicable	76
Bernoulli ---	65
expected value of a count from observed count for Bernoulli ---	75
expected value of a count from observed value for a Poisson ---	69
expected value of Bernoulli ---	65
function	31-33
functions of counting results	64
Gaussian ---	67
normal ---	67
normalized --- function	32
of expected values for a given observed count for a large sampling	69
of expected values for a given observed count for a small sampling	68
of intervals between output pulses from a scale-of-m circuit	117

	<u>Page No.</u>
Distribution:	
Poisson ---	67
Poisson --- not valid	70
renormalized probability --- of background count	76
standard deviation of normal ---	68
standard deviation of Bernoulli ---	66
standard deviation of Poisson ---	67
variance of a total ---	42
variance of expected value of a count for a Poisson ---	69
variance of expected value of a count for Bernoulli ---	75
variance of the --- formed by a sample	42
Divided sample:	
measuring dead time with a ---	114
Efficiency:	4
counting ---	64
detector ---	9
measurement of absolute counting --- by coincidence techniques	127
Electronic:	
estimation of dead time by --- methods	116
scaling circuit	117
Element:	
cofactor of a matrix ---	29
of the matrix	27
Elements:	
diagonal --- of covariance matrix are variances	87
Energy:	
cutoff	87
particle --- spectra	124
Equation:	
inverse of matrix of the normal ---	52
inverse matrix of the normal ---	61
number of terms in the ---	60

	<u>Page No.</u>
Equations:	
algebra of linear ---	26
assumptions underlying decay	3
normal ---	49, 52, 89
normal --- in matrix notation	49
simultaneous linear ---	26
weighted --- for data in matrix form	61, 62
Equilibrium:	
buildup toward ---	8
Equipment:	
additive variance due to counting	86
Estimate:	
best ---	26
best --- of the true value for an experiment	96
graphical --- of the half-life	6
of the half-lives	5
to correct for variances not included in least-squares weighting	56
unbiased --- of variance	42
unbiased --- of variance in least-squares solution	56
unbiased --- of variance of points in least-squares solution	53
Estimated:	
measure of the precision of an individual measurement	42
measure of the precision of the average of a number of measurements	42
Estimation:	
of dead time by electronic methods	116
of dead time by graphical methods	116
of instrumental variance	88
of population variance from a sample	41
of variance due to instrumental variation	87
precision of --- of a mean	40
Evaluating:	
data	92
iteration methods for --- lambda	107
method of --- a decay constant	101

	<u>Page No.</u>
Evaluation:	
graphical --- of lambda	102
of lambda for known multicomponent systems	107
Event rate:	
variance of coincident	123
Events:	
chance coincident ---	120
coincident ---	124
effect of coincident --- on discriminated counts	124
measurement of coincident ---	120
probability of observing N ---	66
probability of recording exactly N ---	65
truly coincident	120
Expansion by minors	29
Expected value:	
algebra of ---	36
distribution of --- for a given observed count for a large sampling	69
distribution of --- for a given observed count for a small sampling	68
of a count from observed count for Bernoulli distribution	75
of a count from observed value for a Poisson distribution	69
of a matrix or vector	52
of any function of x	32
of Bernoulli distribution	65
of n and its variance where we have an actual observed value	68
of net count	77
of the false coincidence rate	123
of x	32
variance of --- of net count	77
variance of --- of a count for a Bernoulli distribution	75
variance of --- of a count for a Poisson distribution	69

Experiment:	
best estimate of the true value for an ---	96
information to be obtained in an --- on radioactive material	4
Extrapolated:	
variance of an --- count rate	79
Extrapolating:	
counts per unit of time	81
Extrapolation:	
errors of ---	79,82
factor	80,81
variance	82
variance of the --- factor	81
F-test for number of components	60
Figure:	
loss	61
of merit	9
Figures:	
significant	61
Fit:	
causes of poor ---	60
chi-squared test of goodness of ---	60
effect of varying an input lambda on the --- of the data to the calculated curve	91
goodness of ---	94
Fluctuations:	
variance due to statistical ---	96
Flux:	
.	5
constant	8
reactor --- measurement	83
varying	8
Form:	
improper --- for function being fitted	94
Formation:	
independent ---	21
net rate of ---	7
of active species	8-10
of active species by decay of a radioactive precursor	10
of species in nuclear reactions	5

	<u>Page No.</u>
Formed:	
analytical scheme when both parent and daughter are --- directly by nuclear reactions	17
variance of the distribution --- by a sample	42
Formulas:	
mathematical techniques for deriving --- for decay of complex systems	21-24
Freedom:	
degrees	43
Function:	
distribution	31-33
improper form for --- being fitted	94
expected value of any --- of x	32
normalized distribution ---	32
Functions:	
distribution --- of counting results	64
Fundamental:	
law of radioactivity	103
law of radioactive decay	3
Gamma spectrum:	
difficulty in interpreting sum peaks in ---	126
Gaussian distribution	67
Geometry:	
of counting arrangement	65
Goodness of fit:	94
chi-squared test of ---	60
Graphical:	
estimate of the half-life	6
estimation of dead time by --- methods	116
evaluation of lambda	102
resolution of decay curve	6
Growth:	
factor	12
factor in branching decay	20
of activity	13
Half-life:	
alteration in the --- value	4,6 91
analytical schemes when --- of parent is much longer than that of daughter	15

Half-life:

by direct measurement of specific activity	103
consecutive decay for parent and daughter with equal ---	11
effect of an impurity on the measurement of a ---	101
least-squares solution for ---	102
measurement of a --- in a single-component system of short or moderate ---	102
measurement of --- in a single-component system with long ---	103
measurement of --- by specific activity	103
measuring dead time using a short-lived species of well-known ---	115
moderate ---	102
precisely known	103

Half-lives:

correction for presence of daughter in a parent-daughter systems where both components have comparable ---	13
difficulty of resolving two species with similar ---	106
estimation of the ---	5
measurement of ---	101
partial ---	20
sensitivity of x to the input ---	95

Half-period	4
-----------------------	---

High count rates	113
----------------------------	-----

High-flux reactors	9
------------------------------	---

Human error	92
-----------------------	----

Identification of components	101
----------------------------------------	-----

Imprecision:

statistical --- of a single radiochemical analysis	95
--------------------------------------------------------------	----

Improper:

form for function being fitted	94
------------------------------------------	----

use of least-squares analysis	94
-----------------------------------------	----

Improving:

values of the input λ	92
-----------------------------------------	----

Impurities	101
----------------------	-----

Impurity:

effect of an --- on the measurement of a half-life	101
--------------------------------------------------------------	-----

	<u>Page No.</u>
Independent formation	21
Individual:	
estimated measure of the precision of an --- measurement	42
quantities which are to be averaged are not uncorrelated .	45
variance of an --- measurement	40
Information:	
loss of --- due to rounding errors	61
to be obtained in an experiment on radioactive material .	47
Instrument malfunction	92
Instrumental variance:	
estimation of	88
magnitude of ---	88
Instrumental variations:	
estimation of variance due to ---	87
Integral:	
treatment of unknown system by --- transforms	110
Intensity:	
sum peak	126
Interpolation:	
variance due to neglect of --- lights	85, 86
Interpreting:	
difficulty in --- sum peaks in gamma spectra	126
Interval:	65
average --- between pulses	117
n-percent confidence ---	33
Intervals:	
distribution of --- between output pulses from a scale-of-m circuit	117
Inverse:	
matrix	29, 30
matrix of the normal equation	61
of matrix of the normal equation	52
Inversion:	
matrix --- by orthogonalizing	61
matrix --- method for resolution of decay curve	87

Inverting:		
a diagonal matrix	63
Ionization:		
measurement of half-life using two --- chambers	104
Iteration:		
methods for evaluating lambda	107
Lambda:		
effect of varying an input --- on the fit of the data to the calculated curve	91
evaluation of --- for known multicomponent systems	107
graphical evaluation of ---	102
improving the values of the input ---	92
iteration methods for evaluating ---	107
measurement of ---	101
Lambdas:		
effect of a slight change in one of the --- resolution	89
systematic effect on resolution caused by slight variations in the values of the ---	84
• Law:		
fundamental --- of radioactive decay	3
fundamental --- of radioactivity	103
Least-squares:		
computational method	61
effect on --- solutions of any unanticipated variances	83
estimate to correct for variances not included in --- weighting	56
improper use of --- analysis	94
precision of each parameter in a --- solution	56
proper weighting factor for use in a --- calculation of decay data	87
proper weighting for ---	50
proper weighting of decay data for a --- solution	83
Least-squares solution:	47
and weighting of data	84
for half-life	99

	<u>Page No.</u>
Least-squares solution:	
in matrix notation	50
to resolve decay data	83
unbiased estimate of variance in ---	56
unbiased estimate of variance of points in ---	53
variances of results of ---	52
weighting factors in ---	48
Limits:	
confidence	34
Linear:	
algebra of --- equations	26
simultaneous --- equations	26
variance of any --- combination of variates	39
Live time	117
Long-lived:	
self-calibrating method for measuring --- decays	105
two or more --- components	107
Losses:	
avoiding --- in slow circuit components	117
--- due to slow register response	117
reducing ---	121
Magnitude:	
instrumental variance	88
Malfunction:	
instrument	92
Mathematical:	
techniques	25
techniques for deriving formulas for decay of complex systems	21-24
Mathematics:	
of radioactive processes	3
statistical	31
Matrices:	
algebra of ---	26
combination of ---	27
multiplication of ---	28

	<u>Page No.</u>
Matrix:	
algebra	26
cofactor of a --- element	29
covariance	39, 46, 52
determinant of a square ---	29
diagonal ---	29
diagonal elements of covariance --- are variances	87
element of the ---	27
expected value of a --- or vector	52
inverse	29, 30
inverse --- of the normal equation	61
inverse of --- of the normal equation	52
inversion by orthogonalizing	61
inversion method for resolution of decay curve	87
inverting a diagonal ---	63
least-squares solution in --- notation	50
lower triangular ---	29
multiplication	28
nonsingular ---	29
normal equations in --- notation	49
orthogonalization process	62
relationships	30
singular ---	29
square ---	29
symmetric ---	29
transpose of a ---	27
upper triangular ---	29
unit ---	29
weighted equations for data in --- form	61, 62
zero ---	29
Matrix-vector:	
multiplication	27
notation	26

	<u>Page No.</u>
Maximum:	
counting rate	113
production of a radioactive species	10
Measure:	
estimated --- of the precision of an individual measure- ment	42
estimated --- of the precision of the average of a number of measurements	42
Measured:	
probability that a --- value of x will fall between A and B .	32
Measurement:	
effect of an impurity on the --- of a half-life	101
estimated measure of the precision of an individual --- .	42
half-life by direct --- of specific activity	103
of a half-life in a single-component system of short or moderate half-life	102
of absolute counting efficiency by coincidence techniques .	127
of coincident events	120
of counter dead time with an oscilloscope	118
of dead time	114
of half-lives	101
of half-life by specific activity	103
of half-life in a single-component system with long half- life	103
of half-life using two ionization chambers	104
of lambda	101
of true disintegration rate	127
reaction cross section ---	83
reactor flux ---	83
variance of an individual ---	40
Measurements:	
averaging --- of unlike variances	45
estimated measure of the precision of the average of a number of ---	42
statistical reliability of --- on nuclear reactions	82

	<u>Page No.</u>
Measuring:	
counter dead time	114
dead time using a short-lived species of well-known half-life	115
dead time with a divided sample	114
Mechanical registers	117
Merit:	
figure of ---	9
Minors:	
expansion by ---	29
moderate half-life	102
Multicomponent:	
detection of --- systems	105
evaluation of lambda for known --- systems	107
unknown --- systems	109
Multiplication:	
matrix ---	28
matrix-vector	27
of matrices	28
N-percent of confidence interval	33
Neglect:	
variance due to --- of interpolation lights	85, 86
Nonindependent: averaging of --- quantities	45
Nonrandom background	76
Nonsingular matrix	29
Normal:	
distribution	67
equations	49, 52, 89
equations in matrix notation	49
inverse matrix of the --- equation	61
inverse of matrix of the --- equation	52
standard deviation of --- distribution	68
Normalized distribution function	32

	<u>Page No.</u>
Notation:	
least-squares solution in matrix ---	5
matrix-vector ---	26
normal equations in matrix ---	49
Nuclear reaction:	
rate of the ---	5
Nuclear reactions:	
formation of species in ---	5
statistical reliability of measurements on ---	82
use of daughter activities to measure ---	13
Number:	
F-test for --- of components	60
of components	101
of standard deviations	92
of terms in the equation	60
variance of best average value for a given --- of replicates	98
Observed count:	
distribution of expected values for a given --- for a large sampling	69
distribution of expected values for a given --- for a small sampling	68
expected value of a count from --- for a Poisson distribution	69
expected value of a count from --- for Bernoulli distribution	75
expected value of n and its variance where we have an actual --- of n	68
variance of ---	84
Observed counting rate:	
effect of dead time on ---	112
Observing:	
decay over long periods of time	103
probability of --- n events	66
Orthogonal vectors	30
Orthogonalization:	
matrix --- process	62
Orthogonalizing:	
matrix inversion by ---	61

Oscilloscope:			
measurement of counter dead time with an ---	118	
Output:			
data	94	
distribution of intervals between --- pulses from a scale-	of-m circuit	117	
Overflow:			
avoidance of computer ---	94	
Parameter:			
precision of each --- in a least-squares solution	56	
Parameters:			
effect of input --- on resolution	89	
Parent:			
analytical scheme when both --- and daughter are formed	directly by nuclear reactions	17	
analytical scheme when --- is much shorter-lived than	daughter	16	
analytical schemes when half-life of --- is much longer	than that of daughter	15	
consecutive decay for --- and daughter with equal half-	life	11	
Parent-daughter:			
correction for presence of daughter in --- systems	where both components have comparable half-lives	13	
correction for presence of daughter in --- systems	where daughter is considerably longer-lived than	parent	13
decay of a --- pair where parent is the longer-lived	species	14	
decay of a --- pair where parent is the shorter-lived	species	14	
decay of a --- system	12	
Partial:			
decay constants	19	
half-lives	20	
Particle energy spectra	124	
Pattern:			
systematic --- in residuals	94	

	<u>Page No.</u>
Percentage error	92
Period	4
Periods:	
observing decay over long --- of time	103
Photopeaks:	
counting single ---	87
Points:	
suspect	93, 94
unbiased estimate of variance of --- in least-squares solution	53
Poisson distribution:	67
background correction when --- is not applicable	76
distribution not valid	70
expected value of a count from observed value for a ---	69
standard deviation of ---	67
variance of expected value of a count for a ---	69
Poor fit:	
causes of	60
Population:	
estimation of --- variance from a sample	41
Precision:	33, 94
estimated measure of the --- of an individual measure- ment	42
estimated measure of the --- of the average of a number of measurements	42
of absolute counting by coincidence	127
of each parameter in a least-squares solution	56
of estimation of a mean	40
of the amount of each component present	87
Precursor:	
formation of active species by decay of a radioactive ---	10
Probability:	13
of disintegrating and producing a count	64
of observing n events	66
of recording exactly n events	65

Probability:	
that a measured value of x will fall between A and B	32
that precisely n counts will be recorded during the counting interval	65
renormalized --- distribution of background count	76
Product:	
destruction of	10
variance of a	37
Production:	
maximum --- of a radioactive species	10
Pulses:	
average interval between ---	117
distribution of intervals between out --- from a scale- of-m-circuit	117
Quotient:	
variance of a ---	38
Radiation scattering and absorbing	65
Radioactive:	
formation of active species by decay of a --- precursor	10
fundamental law of --- decay	3
information to be obtained in an experiment on --- material	4
mathematics of --- processes	3
maximum production of a --- species	10
Radioactivity:	
fundamental law of ---	103
Radiochemical:	
statistical imprecision of a single --- analysis	95
Random errors	25, 31
Reaction:	
rate of the nuclear ---	5
cross section measurement	83
Reactions:	
analytical scheme when both parent and daughter are formed directly by nuclear ---	17
formation of species in nuclear ---	5

	<u>Page No.</u>
Reactions:	
statistical reliability of measurements on nuclear --- .	82
use of daughter activities to measure nuclear --- . .	13
Reactor:	
flux measurement	83
Reactors, high-flux	9
Reducing, losses	121
Reduction in rounding error	63
Register:	
losses due to slow --- response	117
Registers:	
mechanical	117
Regularizing:	
action of a scaling circuit	117
action of scaling circuits	121
Rejecting:	
additional criterion for --- data	93
Rejection:	
criteria for --- of data	92
of data	84
simple scheme for data ---	93
Reliability:	
statistical --- of measurements on nuclear reactions .	82
Renormalized probability distribution of background count . .	76
Replicate:	
analysis of --- samples	35, 95
averaging --- values	98
total variance for each ---	96
variance between --- samples	96
weighting factor to average the --- results	96
Replicates:	95
best average value for a given number of ---	98
variance of best average value for a given number of --- .	98

	<u>Page No.</u>
Residual:	
weighted	92
Residuals:	
systematic pattern in ---	94
weighted residuals of the system	63
Resolution:	
effect of a slight change in one of the lambdas on ---	89
effect of input parameters on ---	89
graphical --- of decay curve	6
matrix inversion method for --- of decay curve	87
of decay data	83
systematic effect on --- caused by slight variations in the values of the lambdas	84
Resolve:	
least-squares solution to --- decay data	83
Resolving:	
difficulty of --- two species with similar half-lives	106
technique for --- data	5
the decay data	4
time	112, 120
Response:	
losses due to slow register ---	117
Rounding:	
error	61
loss of information due to --- errors	61
reduction in --- error	63
Row vector	27
Sample:	
estimation of population variance from a ---	41
measuring dead time with a divided ---	114
total counting rate of a ---	11
variance of the distribution formed by a	42

	<u>Page No.</u>
Samples:	
analysis of replicate ---	35, 95
split	114
variance between replicate ---	96
Sampling:	68
averaging and --- statistics	39
distribution of expected values for a given observed count for a large ---	69
Saturation:	7
of a counter	113
Scale-of-m:	
distribution of intervals between output pulses from a --- circuit	117
Scaling:	
advantage to be gained by use of --- factors	120
circuits	85, 117
electronic --- circuit	117
regularizing action of a --- circuit	117
regularizing action of --- circuits	121
Scattering:	
radiation --- and absorbing	65
Self-calibrating method for measuring long-lived decays	105
Sensitivity of x to the input half-lives	95
Significant figures	61
Simultaneous linear equations	26
Single:	
statistical imprecision of a --- radiochemical analysis	95
Single-component:	
measurement of a half-life in a --- system with long half-life	103
measurement of a half-life in a --- system of short or moderate half-life	102
Singles:	
ratio of sums to ---	124, 125

	<u>Page No.</u>
Singular matrix	29
Solid:	
coincidence counting of large --- angles	122
Solution:	
least-squares --- and weighting of data	84
least-squares --- for half-life	102
least-squares --- in matrix notation	50
least-squares ---	47
least-squares --- to resolve decay data	83
precision of each parameter in a least squares ---	56
proper weighting of decay data for a least squares ---	83
unbiased estimate of variance in least squares ---	56
unbiased estimate of variance of points in least squares ---	53
Solutions:	
effect on least squares --- of any unanticipated variances	83
variances of results of least squares ---	52
weighting factors in least squares ---	48
Species:	
difficulty of resolving two --- with similar half-lives	106
formation of active ---	8-10
formation of active --- by decay of a radioactive precursor	10
formation of --- in nuclear reactions	5
maximum production of a radioactive ---	10
measuring dead time using a short-lived --- of well-known half-life	15
Specific activity:	
half-life by direct measurement of ---	103
measurement of half-life by ---	103
Spectra:	
difficulty in interpreting sum peaks in gamma ---	126
particle energy ---	124
Split samples	114

	<u>Page No.</u>
Square matrix:	29
determinant of a ---	29
Squares:	
weighted sum of the ---	49
Standard deviation:	
of Bernoulli distribution	66
of normal distribution	68
number of ---	92
of Poisson distribution	67
of x	33
Statistical:	
imprecision of a single radiochemical analysis	95
mathematics	31
reliability of measurements on nuclear reactions	82
variance due to --- fluctuations	96
Statistics:	
averaging and sampling ---	39
of counting data	64
substitution of a spurious zero-time	94
Sum:	
of energies	124
variance of a ---	37
weighted --- of the squares	49
Sum peak:	124, 125, 127
decay rate for the false ---	127
difficulty in interpreting --- in gamma spectra	126
intensity of the ---	126
Sums:	
false ---	125
ratio of --- to singles	124, 125
Suspect point	93, 94
Symmetric matrix	29

Systematic:

effect on resolution caused by slight variations in the values of the lambdas	84
errors	95
pattern in residuals	94

Target:

material depleted	9
-----------------------------	---

Technique:

for resolving data	5
variance due to analytical ---	96

Techniques:

mathematical	25
mathematical --- for deriving formulas for decay of complex systems	21, 22 23, 24
measurement of absolute counting efficiency by coincidence ---	127

Test:

chi-square	88
chi-square --- of goodness of fit	60

Time:

actual correction for dead ---	113
dead	110, 112, 113
effect of dead --- on observed counting rate	112
estimation of dead --- by electronic methods	116
estimation of dead --- by graphical methods	116
extrapolating counts per unit ---	81
live ---	117
measurement of counter dead --- with an oscilloscope	118
measurement of dead ---	114
measuring counter dead ---	114
measuring dead --- using a short-lived species of well-known half-life	115
measuring dead --- with a divided sample	114
observing decay over long periods of ---	103
resolving ---	112, 120

	<u>Page No.</u>
Transforms:	
treatment of unknown system by integral ---	110
Transition:	
cross over	125
Transpose:	
of a matrix	27
of a vector	27
Triangular:	
lower --- matrix	29
upper --- matrix	29
Truly coincident events	120
Two-component:	
decay curve	6
decay curve of a --- system	106
Typical errors	92
Unanticipated:	
effect on least-squares solutions of any --- variances	83
Unbiased estimate:	
of variance	42
of variance in least-squares solution	56
of variance of points in least-squares solution	53
Uncorrelated:	
individual quantities which are to be averaged are not ---	45
variates	35
Unknown:	
comparison of the decay of the --- species with that of a species with precisely known half-life	103
multicomponent systems	109
treatment of --- system by integral transforms	110
Unlike variances:	
averaging measurements of ---	45
Valid:	
Poisson distribution not ---	7

	<u>Page No.</u>
Values:	
algebra of expected ---	36
averaging replicate ---	98
distribution of expected --- for a given observed count for a large sample	68
distribution of expected --- for a given observed count for a small sample	69
improving the --- of the input lambda	92
systematic effect on resolution caused by slight variations in the --- of the lambdas	84
Variance:	
additive --- due to counting equipment	86
analysis of ---	35, 95
average having the least ---	45
between replicate samples	96
due to analytical technique	96
due to counter variation	87
due to neglect of interpolation lights	85, 86
due to statistical fluctuations	96
estimation of instrumental ---	88
estimation of population --- from a sample	41
estimation of --- due to instrumental variation	87
expected value of n and its --- where we have an actual observed value of n	68
extrapolation ---	82
magnitude of instrumental variance	88
of a difference	37
of a net count	76
of a product	37
of a quotient	38
of a sum	37
of a total distribution	42
of a weighted average	47
of an average	40
of an extrapolated count rate	79

	<u>Page No.</u>
Variance:	
of an individual measurement	40
of an observed count	84
of any linear combination of variates	39
of best average value for a given number of replicates .	98
of coincident event rate	123
of expected value of a count for a Poisson distribution .	69
of expected value of a count for Bernoulli distribution .	75
of expected value of net count	77
of ln N	102
of net counting rate	84
of the distribution formed by a sample	42
of the extrapolation factor	81
of the false coincident rate	123
of x	33
total --- for each replicate	96
unbiased estimate of ---	42
unbiased estimate of --- in least-squares solution . .	56
unbiased estimate of --- of points in least-squares solution	53
Variations:	
algebra of ---	36
averaging measurements of unlike ---	45
diagonal elements of covariance matrix are ---	87
effect on least-squares solutions of any unanticipated ---	83
estimate to correct for variances not included in least- squares weighting	56
of results of least-squares solutions	52
which contribute to the weighting factor	83
Variate	31
Variates:	
uncorrelated	35
variance of any linear combination of ---	39

	<u>Page No.</u>
Variation:	
estimation of variance due to instrumental ---	87
variance due to counter ---	87
Variations:	
analytical	95
systematic effect on resolution caused by light --- in the values of the lambdas	84
Varying:	
flux	8
effect of --- an input lambda on the fit of the data to the calculated curve	91
Vector:	
column	27
row	27
expected value of a matrix or ---	52
transpose of a ---	27
Vectors:	
orthogonal ---	30
Weight:	
in the averaging process	44
Weighted:	
average	44, 45
equations for data in matrix form	61, 62
residual	92
residuals of the system	63
sum of the squares	49
variance of a --- average	47
Weighting:	
estimate to correct for variances not included in least- squares ---	56
factor to average the replicate results	96
factors	45
factors in least-squares solutions	48
least-squares solution and --- of data	84

Weighting:

proper --- factors for use in a least-squares calculation of decay data	87
proper --- for least-squares	50
proper --- of decay data for a least-squares solution .	83
variances which contribute to the --- factor	83

Zero-time:

substitution of a spurious ---	94
variance of $\ln N$	102

MONOGRAPHS IN THE FIELD OF CHEMISTRY AND PHYSICS
OF THE NATIONAL BUREAU OF STANDARDS

Copies of the following monographs are available from the Clearinghouse for Federal Scientific and Technical Information, National Bureau of Standards, U. S. Department of Commerce, Springfield, Va. 22151

- Aluminum and Gallium, NAS-NS-3032, \$0.50
Americium and Curium, NAS-NS-3006, \$0.75
Antimony, NAS-NS-3033, \$0.50
Arsenic, NAS-NS-3002, (Rev.) 1965 \$1.00
Astatine, NAS-NS-3012, \$0.50
Barium, Calcium, and Strontium, NAS-NS-3010, \$1.25
Beryllium, NAS-NS-3013, \$0.75
Cadmium, NAS-NS-3001, \$0.75
Carbon, Nitrogen, and Oxygen, NAS-NS-3019, \$0.50
Cesium, NAS-NS-3035, \$0.75
Chromium, NAS-NS-3007, (Rev.) 1964 \$0.75
Cobalt, NAS-NS-3041, \$1.00
Copper, NAS-NS-3027, \$0.75
Fluorine, Chlorine, Bromine, and Iodine, NAS-NS-3005, \$0.50
Francium, NAS-NS-3003, \$0.50
Germanium, NAS-NS-3043, \$0.50
Gold, NAS-NS-3036, \$0.50
Indium, NAS-NS-3014, \$0.50
Iridium, NAS-NS-3045, \$0.50
Iron, NAS-NS-3017, \$0.50
Lead, NAS-NS-3040, \$1.75
Magnesium, NAS-NS-3024, \$0.50
Manganese, NAS-NS-3018, \$0.50
Mercury, NAS-NS-3026, \$0.50
Molybdenum, NAS-NS-3009, \$0.50
Nickel, NAS-NS-3051, \$0.50
Niobium and Tantalum, NAS-NS-3039, \$0.75
Osmium, NAS-NS-3046, \$0.50
Palladium, NAS-NS-3052, \$0.75
Phosphorus, NAS-NS-3056, \$0.50
Platinum, NAS-NS-3044, \$0.50
Polonium, NAS-NS-3037, \$0.75
Potassium, NAS-NS-3048, \$0.50
Protactinium, NAS-NS-3016, \$1.00
Radium, NAS-NS-3057, \$2.25
Rare Earths—Scandium, Yttrium, and Actinium, NAS-NS-3020, \$3.00
Rare Gases, NAS-NS-3025, \$0.75
Rhenium, NAS-NS-3023, \$0.50
Rhodium, NAS-NS-3003, (Rev.) 1965 \$1.00
Rubidium, NAS-NS-3053, \$0.50
Ruthenium, NAS-NS-3029, \$1.00
Selenium, NAS-NS-3030, (Rev.) 1965 \$1.00
Silicon, NAS-NS-3049, \$0.50
Silver, NAS-NS-3047, \$0.75
Sodium, NAS-NS-3055, \$0.50
Sulfur, NAS-NS-3054, \$0.50
Technetium, NAS-NS-3021, \$0.50
Tellurium, NAS-NS-3038, \$0.50
Thorium, NAS-NS-3004, \$0.75
Tin, NAS-NS-3023, \$0.75
Titanium, NAS-NS-3034, \$0.50
Transcurium Elements, NAS-NS-3031, \$0.50
Tungsten, NAS-NS-3042, \$0.50
Uranium, NAS-NS-3050, \$3.50
Vanadium, NAS-NS-3022, \$0.75
Zinc, NAS-NS-3015, \$0.75
Zirconium and Hafnium, NAS-NS-3011, \$0.50
- Applications of Computers to Nuclear and Radiochemistry, NAS-NS-3107, \$2.50
Application of Distillation Techniques to Radiochemical Separations, NAS-NS-3108, \$0.50
Detection and Measurement of Nuclear Radiation, NAS-NS-3105, \$1.50
Liquid-liquid Extraction with High-molecular-weight Amines, NAS-NS-3101, \$1.00
Low-level Radiochemical Separations, NAS-NS-3103, \$0.50
Paper Chromatographic and Electromigration Techniques in Radiochemistry, NAS-NS-3106, \$0.50
Processing of Counting Data, NAS-NS-3109, \$1.75
Rapid Radiochemical Separations, NAS-NS-3104, \$1.25
Separations by Solvent Extraction with Tri-n-octylphosphine Oxide, NAS-NS-3102, \$0.75