UNCLASSIFIED

OAK RIDGE NATIONAL LABORATORY LIBRARY
3 4456 0022224 8

CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION

ORNL-2134
Progress

cy. 6

MATHEMATICS PANEL

SEMIANNUAL PROGRESS REPORT

FOR PERIOD ENDING JUNE 30, 1956

CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON
If you wish someone else to see this document, send in name with document and the library will arrange a loan.

OAK RIDGE NATIONAL LABORATORY
OPERATED BY
UNION CARBIDE NUCLEAR COMPANY
A Division of Union Carbide and Carbon Corporation
UCC
POST OFFICE BOX P · OAK RIDGE, TENNESSEE

UNCLASSIFIED

Printed in USA. Price 25 cents. Available from the

Office of Technical Services
U. S. Department of Commerce
Washington 25, D. C.

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, express or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission to the extent that such employee or contractor prepares, handles or distributes, or provides access to, any information pursuant to his employment or contract with the Commission.

UNCLASSIFIED

ORNL-2134

Contract No. W-7405-eng-26

MATHEMATICS PANEL

SEMIANNUAL PROGRESS REPORT

for

Period Ending June 30, 1956

A. S. Householder, Chief

Edited by
A. S. Householder

DATE ISSUED

SEP 11 1956

OAK RIDGE NATIONAL LABORATORY
Operated by
UNION CARBIDE NUCLEAR COMPANY
A Division of Union Carbide and Carbon Corporation
Post Office Box P
Oak Ridge, Tennessee

UNCLASSIFIED

OAK RIDGE NATIONAL LABORATORY LIBRARY
3 4456 0022224 8

ORNL-2134
Progress

INTERNAL DISTRIBUTION

1. C. E. Center
2. Biology Library
3. Health Physics Library
4. Reactor Experimental Engineering Library
5–7. Central Research Library
8–14. Laboratory Records Department
15. Laboratory Records, ORNL RC
16. A. M. Weinberg
17. L. B. Emlet (K-25)
18. J. P. Murray (Y-12)
19. J. A. Swartout
20. R. R. Dickison
21. E. D. Shipley
22. E. J. Murphy
23. M. L. Nelson
24. A. S. Householder
25. A. H. Snell
26. E. H. Taylor
27. A. Hollaender
28. W. H. Jordan
29. D. W. Cardwell
30. M. T. Kelley
31. K. Z. Morgan
32. E. M. King
33. C. E. Winters
34. E. O. Wollan
35. E. P. Blizard
36. R. S. Livingston
37. C. P. Keim
38. J. W. Givens
39. C. D. Susano
40. G. C. Williams
41. R. A. Charpie
42. V. C. Klema
43. N. M. Dismuke
44. C. P. Hubbard
45. A. W. Kimball
46. M. R. Arnette
47. D. D. Cowen
48. P. M. Reyling
49. A. C. Downing, Jr.
50. J. Z. Hearon
51. S. G. Campbell
52. C. L. Gerberich
53. J. A. Lane
54. R. W. Johnson
55. M. J. Skinner
56. J. L. Gabbard
57. G. E. Boyd
58. S. J. Cromer
59. E. C. Long
60. C. L. Bradshaw
61. J. P. Kelly (K-25)
62. S. E. Atta
63. ORNL – Y-12 Technical Library, Document Reference Section

EXTERNAL DISTRIBUTION

64. R. F. Bacher, California Institute of Technology
65. Division of Research and Development, AEC, ORO
66–768. Given distribution as shown in TID-4500 under Physics category (300 copies – OTS)

DISTRIBUTION PAGE TO BE REMOVED IF REPORT IS GIVEN PUBLIC DISTRIBUTION

Mathematics Panel progress reports previously issued in this series are as follows:

ORNL-345	December, January, February, 1948–1949
ORNL-408	Period Ending July 31, 1949
ORNL-516	Period Ending October 31, 1949
ORNL-634	Period Ending January 31, 1950
ORNL-726	Period Ending April 30, 1950
ORNL-818	Period Ending July 31, 1950
ORNL-888	Period Ending October 31, 1950
ORNL-979	Period Ending January 31, 1951
ORNL-1029	Period Ending April 30, 1951
ORNL-1091	Period Ending July 31, 1951
ORNL-1151	Period Ending October 31, 1951
ORNL-1232	Period Ending January 31, 1952
ORNL-1290	Period Ending April 30, 1952
ORNL-1360	Period Ending July 31, 1952
ORNL-1435	Period Ending October 31, 1952
ORNL-1498	Period Ending January 31, 1953
ORNL-1588	Period Ending June 30, 1953
ORNL-1662	Period Ending December 31, 1953
ORNL-1751	Period Ending June 30, 1954
ORNL-1842	Period Ending December 31, 1954
ORNL-1928	Period Ending June 30, 1955
ORNL-2037	Period Ending December 31, 1955

CONTENTS

INTRODUCTION 1
The Oracle 1
Lectures 1
Papers Presented at Meetings 2
Publications 2
Oracle Programing Classes 2
Personnel Changes 3
NUMERICAL ANALYSIS RESEARCH 4
Computational Stability in the Solution of Differential Equations 4
Estimation of Mean Survival Time from Samples with Losses 5
Eigenvalues and Vectors for the Product of Two Symmetric Matrices 7
Approximation Derived from the Generalization of Taylor's Expansion 8
Experiments with the Inverse of a Hilbert Matrix 8
Inverse of a 190 × 190 Matrix 9
Cross Section for γ,n Reaction in Be^9 and in C^{13} 9
Theoretical Studies in Numerical Analysis. Lower Bound for Separation of Characteristic Vectors 10
Vector-Block Matrix Codes 10
A Matrix Inversion and Linear-Equation-Solver Program 10
Test for the Effect of Neutron Irradiation on Longevity in Mice 12
A Model for the Multiple-Cause-of-Death Problem 13
A Variance-Stabilizing Transformation for Certain *Paramecium* Experiments 16
Curve-Plotter Code Edit 17
Damage to Thin Slabs of Tissue from Neutron Irradiation 18
Determination of the Breit-Wigner Constants for Xenon 18
Fission-Product Problem 18
Hand Computations 19
Calculation of Stresses in Piping System 19
Cylinder Activation and Solid-Angle Calculations 19
Echellogram Data Reduction 21
Variance Calculations in a *Drosophila* Experiment 22

MATHEMATICS PANEL SEMIANNUAL PROGRESS REPORT

INTRODUCTION

THE ORACLE

From January through June, a total of 1552 hr of computing time was used by programmers for "debugging" and running problems.

Modifications to improve reliability of the Magnetic Tape Auxiliary Memory were completed. Each binary digit is stored simultaneously in two channels separated by half the tape width. This means that each 40-bit machine word is broken down into two 20-bit characters and stored consecutively on the tape. The tape unit is thus a serial-parallel device. Although the packing density remains unchanged, the number of words per block is 128.

LECTURES

ORINS Lectures. – The following lectures were given under the Traveling Lecture Program:

A. C. Downing, *Understanding Digital Computers – An Introduction to Programing*, University of Richmond, January 10, 1956; University of Kentucky, January 12, 1956; Agricultural and Mechanical College of Texas, February 7, 1956; University of Texas, February 9, 1956; North Texas State College, February 10, 1956; North Carolina State College, March 8, 1956. *Some Inverse Characteristic Value Problems*, University of North Carolina, March 7, 1956.

A. W. Kimball, *Errors of the Third Kind in Statistical Consulting*, Virginia Polytechnic Institute, April 27, 1956.

Other Lectures. – The following lectures were also given during the report period:

A. S. Householder, *The Position of the University in the Field of High Speed Computation and Data Handling*, University of Cincinnati, January 21, 1956. *Generated Error in Digital Computation*, The Ramo-Wooldridge Corporation, Los Angeles, February 17, 1956. *Latent Roots of Symmetric Matrices*, February 16, 20, 1956, and *Matrix Norms in Numerical Analysis*, February 21, 1956, University of California. *Opportunities for Mathematicians in Government Sponsored Research*, Alabama College, March 30, 1956. *Matrix Norms in Numerical Analysis*, Wayne University, April 11, 1956. *Solving Problems with Digital Computers*, American Physical Society, Washington, D.C., April 27, 1956. *Applications of Matrix Norms in Numerical Analysis*, University of South Carolina, May 17, 1956. *Error Analysis and Linear Problems*, series of lectures at the Special Summer Session on Digital Computers, University of Michigan, June 20–22, 1956. *Trends in Numerical Analysis*, Society for Engineering Education, Ames, Iowa, June 26, 1956.

A. C. Downing, *On the Convergence of Finite Difference Solutions of the Multiregion Diffusion Problem*, University of Michigan, March 22, 1956. *The Role of Digital Computers in Engineering*, Richmond Professional Institute, January 10, 1956.

J. W. Givens, *The Characteristic Value-Vector Problem*, February 15, 1956, and *The Linear Equations Problem*, February 16, 1956, Louisiana State University (Third Annual High-Speed Computer Conference). *Logical Engines: The New Digital Computers*, University of Tennessee, March 20, 1956 (Sigma Xi Lecture). *The Characteristic Value-Vector Problem*, Wayne University, June 1, 1956 (Mathematics Colloquium).

A. W. Kimball, *Statistical Methods in Radiation Biology*, University of Tennessee, May 8, 1956.

PAPERS PRESENTED AT MEETINGS

Mathematical Association of America, Athens, Georgia, March 15, 16, 1956

N. M. Dismuke, *A Comparison of the Deming and Garwood Fitting Methods.*

A. C. Downing, *On the Behavior of Solutions of Ordinary Finite Difference Equations.*

T. W. Hildebrandt, ORINS Fellow, *From Matrices to Tiles.*

PUBLICATIONS

T. H. Pittenger, A. W. Kimball, and K. C. Atwood, "Control of Nuclear Ratios in *Neurospora* Heterokaryons," *Am. J. Botany* **42**, 954–957 (December 1955).

C. L. Bradshaw, *A Three-Region, Two-Group, Two-Dimensional Reactor Code for the Oracle*, ORNL-2047 (March 27, 1956).

A. C. Downing, *A New Approach to Computer Design*, ORNL CF-56-3-61 (March 20, 1956).

A. S. Householder, "Bibliography on Numerical Analysis," *J. Assoc. Computing Mach.* **3**, 85 (1956).

A. S. Householder, *Some Observations on Reactor Criticality Calculations*, ORNL CF-56-4-191 (May 10, 1956).

A. S. Householder, "Presidential Address to the ACM, Philadelphia, September 14, 1955," *J. Assoc. Computing Mach.* **3**, 1 (1956).

A. S. Householder, "The Position of the University in the Field of High Speed Computation and Data Handling," *Computers and Automation* **5**(5), 6 (1956).

ORACLE PROGRAMING CLASSES

Three Oracle programing classes were given during the period from January through June. For each class the lectures include 10 to 15 hr of basic machine operation and programing techniques, 5 hr of problem discussion, and 5 hr of special lecture dealing

with more advanced programing techniques. The breakdown of the total group by division is given below:

Applied Nuclear Physics	12
Chemistry	11
Mathematics Panel	7
Electronuclear Research	6
Health Physics	5
Reactor Experimental Engineering	4
Instrumentation and Controls	3
Operations	3
Oak Ridge School of Reactor Technology	3
Stable Isotopes	3
Directors	2
Physics	2
Aircraft Reactor Engineering	1
Chemical Technology	1
Materials Chemistry	1
Metallurgy	1
Solid State	1

The above list was obtained by taking the roll on the sixth day of meeting. Of this total of 66 people, 13 are not permanently employed at ORNL.

PERSONNEL CHANGES

The following are new employees in the Mathematics Panel: M. A. Kastenbaum, D. A. Gardiner, F. J. Witt, and E. Leach. G. W. Medlin and M. C. McCord are on temporary assignment to the Panel. T. W. Hildebrandt and R. G. Cornell, ORINS Fellows, completed their theses and received their doctoral degrees. T. W. Hildebrandt, an employee of the ORINS University Relations Division, is assigned to the Mathematics Panel in connection with the Oracle Applications Program sponsored by ORNL and ORINS. Hildebrandt will work directly with the University Research personnel in preparing problems for Oracle computation.

There was one termination: V. C. Klema.

NUMERICAL ANALYSIS RESEARCH

COMPUTATIONAL STABILITY IN THE SOLUTION OF DIFFERENTIAL EQUATIONS

Participating Member of Panel: A. S. Householder

The matrix norms have been used effectively in the study of this problem in the case of linear differential equations, partial and ordinary. Extending the application of them to nonlinear equations and to the examination of truncation error seems to be possible in at least some cases but has not yet been carried out.

The case considered here is that of partial differential equations in a rectangular region with constant coefficients. Some cases of variable coefficients have been considered but will not be discussed here. Given any grid with uniform spacing, the computational problem becomes that of solving a system of linear algebraic equations of the form

$$Ax = b \quad , \tag{1}$$

where the matrix A depends primarily upon the nature of the finite difference approximation scheme and the vector b depends upon the boundary values. In the explicit or "marching" schemes the matrix A is triangular, and in all cases it can be partitioned into a form such as

$$A = \begin{pmatrix} B & C' & \cdots \\ C & B & \cdots \\ \cdot & \cdot & \cdot \end{pmatrix} ,$$

where (1) the order of each block is equal to the number of grid points in a horizontal row; (2) the number of blocks in any row (or column) is equal to the number of grid points in a vertical row; (3) blocks along the diagonal are identical and those along any codiagonal are identical; and (4) nonnull elements occur only along the diagonal blocks and a few (generally one or two) codiagonals.

The problem of computational stability can be phrased as follows: Suppose that, by any method, a vector x^* (of digital elements) has been obtained and is assumed to represent an approximate solution to system 1; then let

$$A(x - x^*) = d \quad .$$

Although the vector d is not directly available because of rounding errors, the following decomposition can be made:

$$A(x - x^*) = [b - (Ax^*)^*] + [(Ax^*)^* - Ax^*] \quad ,$$

where (Ax^*) represents the machine product of A by x^*. Hence the first vector on the right is directly available, and the second has elements that can at least be bounded. Consequently, given any norm, the norm $||d||$ of d can be bounded. Since

$$x - x^* = A^{-1}d \quad ,$$

$$||x - x^*|| \leqq ||A^{-1}|| \cdot ||d|| \ . \tag{2}$$

The order of the matrix A is, of course, equal to the total number of grid points and therefore the norm is a function of n and m, the number of points in the two directions. The primary interest lies in the rate of increase of this function with n and m.

A number of methods have been examined for parabolic, hyperbolic, and elliptic equations. Since the submatrices B, C, ... are of quite simple structure, their proper values are readily obtained, and the norm $||A^{-1}||_e$ is easily expressible in terms of the norms of these submatrices. In fact, the submatrices are generally expressible as polynomials in either $J + J^T$ or $J - J^T$, where

$$J = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} .$$

Also, the partitioned form of A^{-1} often involves expressions of the form $\psi(K)$ or $\omega(K)$, where K is a submatrix of A and where ψ and ω are the characteristic polynomials of $J + J^T$ and $J - J^T$, respectively.

An important case of variable coefficients arises in the two-dimensional criticality problem. For each group the matrix of the equations has the form

$$A = \begin{pmatrix} A_1 & -C_1 & 0 & \cdots \\ -B_2 & A_2 & -C_2 & \cdots \\ 0 & -B_3 & A_3 & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} ,$$

which can be factored with matrices of the form

$$A = \begin{pmatrix} D_1 & 0 & 0 & \cdots \\ -B_2 & D_2 & 0 & \cdots \\ 0 & -B_3 & D_3 & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} I & -\Gamma_1 & 0 & \cdots \\ 0 & I & -\Gamma_2 & \cdots \\ 0 & 0 & I & \cdots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} .$$

The two factors have readily determined inverses with norms that are readily bounded in many cases.

ESTIMATION OF MEAN SURVIVAL TIME FROM SAMPLES WITH LOSSES

Origin: W. L. Russell, Biology Division

Participating Members of Panel: A. W. Kimball, G. J. Atta

Background and Status. – The F_1 progeny from two groups of male mice, one x irradiated with 600 r and one control, had been set aside for a special genetics study. From

time to time, progeny in both groups were removed for study and were later sacrificed. Out of 183 progeny from the control group and 138 progeny from the irradiated group, 42 and 52, respectively, were never used in the special study and were kept until they died natural deaths. It is desired now to view this experiment as a longevity study in which some of the original mice were, from time to time, lost from observation, that is, removed and intentionally sacrificed.

The two groups of 42 and 52 progeny cannot be regarded as a random sample from the population, and so conventional statistical methods break down. However, a new method proposed by E. L. Kaplan and P. Meier (*Nonparametric Estimation of Survivorship*, unpublished) provides a solution to the problem. In the context of the present problem, the natural death of an animal is called a "death," and the removal and subsequent sacrifice of an animal is called a "loss." In a sample containing originally N animals, the first step in the procedure is to rank the times t_i $(i = 1, \ldots, N)$ of death or loss in order of increasing magnitude; that is,

$$t_1 \leqq t_2 \leqq \cdots \leqq t_N .$$

From this ordered set can be constructed an ordered subset,

$$t_{r_1} \leqq t_{r_2} \leqq \cdots \leqq t_{r_n} ,$$

composed of the n times at death, where $1 \leqq r_j \leqq N$ for all j and $r_1 < r_2 < \ldots < r_n$. From this subset the quantities

$$P\left(t_{r_j}\right) = \prod_{k=1}^{j} \left(\frac{N - r_k}{N - r_k + 1}\right)$$

and

$$R_j = P\left(t_{r_j}\right) \cdot \left(t_{r_{j+1}} - t_{r_j}\right)$$

are computed for each j $(j = 1, \ldots, n)$ (ref 1). Then an estimate of the mean survival time, μ, is given by

$$\hat{\mu} = t_{r_1} + \sum_{j=1}^{n} R_j ,$$

and an estimate of the variance of $\hat{\mu}$ is given by

$$\hat{V}(\hat{\mu}) = \frac{n}{n+1} \sum_{k=1}^{n} \frac{A_k^2}{(N - r_k)(N - r_k + 1)} ,$$

[1]If $r_n = N$, $P\left(t_{r_n}\right) = 0$, and $t_{r_{n+1}}$ need not be defined. If $r_n < N$, $t_{r_{n+1}}$ should be taken as equal to $t_{r_{n+1}}$.

where

$$A_k = \sum_{j=k}^{n} R_j \ .$$

If the last observation in the sample is a death, $\hat{\mu}$ is an unbiased estimate of μ. If the last observation is a loss, $\hat{\mu}$ may be regarded as an estimate of the mean life in the population truncated at the time of the first observed loss following the last observed death.

In the present example both samples were such that the last observations were deaths. Based on the method just described, estimates were found to be as follows:

	Survival (days)	
	Controls	Irradiated
$\hat{\mu}$	887.5	855.1
$\sqrt{\hat{V}(\hat{\mu})}$	18.9	19.9

Since the observed difference is only about one standard error, this experiment cannot be considered as evidence for an x-radiation effect on longevity in mice. It must be remembered, however, that the effective sample sizes are 42 and 52, probably not sufficiently large for a sensitive test of significance.

EIGENVALUES AND VECTORS FOR THE PRODUCT OF TWO SYMMETRIC MATRICES

Origin: F. A. Matsen, University of Texas

Participating Members of Panel: M. R. Arnette, A. C. Downing

Background and Status. – The code for finding eigenvalues and vectors for the product of two symmetric matrices was run on the Oracle for two sets of matrices, each of order 20.

Origin: P. C. Sharrah, University of Arkansas

Participating Member of Panel: M. R. Arnette

Background and Status. – The same Oracle code was run for four sets of matrices, each of order 12.

APPROXIMATION DERIVED FROM THE GENERALIZATION OF TAYLOR'S EXPANSION

Participating Member of Panel: S. E. Atta

Finite expansions for certain functions have been developed from the generalization of Taylor's expansion.[2] The fundamental expansion is

$$(1) \quad f(x) = f(a) + \sum_{k=1} \frac{(m+n-k)!}{(m+n)!} \left[{}_mC_k f^k(a) - (-1)^k \, {}_mC_k f^k(x) \right] (x-a)^k + R \ ,$$

where

$$R = (-1)^n \frac{m!\, n!\, (x-a)^{m+n+1}}{(m+n)!\,(m+n+1)!} \, f^{(m+n+1)}(\theta) \ , \qquad a \leqq \theta \leqq x \ .$$

By letting $m = n$, rapid convergence can be obtained for many functions. It is interesting to develop expansions from Eq. 1 for certain functions and to compare them with the familiar Taylor expansions. Approximations were developed for e^x, log nx, powers, and roots. The expansion

$$(2) \qquad e^x = \frac{1 + \sum_{k=1}^{n} C_k x^k}{1 + \sum_{k=1}^{n} (-1)^k C_k x^k} \ , \qquad -1 \leqq x \leqq 1 \ ,$$

where

$$C_{k+1} = \frac{n-k}{(2n-k)(k+1)} C_k \ , \qquad C_1 = \tfrac{1}{2} \ ,$$

was programed for the Oracle in scaled arithmetic. The e^x can be computed to 12 decimals by taking $n = 6$, whereas the familiar Taylor finite expansion requires $n = 17$ for a comparable accuracy. The expansion for log nx has not been programed. However, an estimation of the remainder term indicates that this expansion will converge only twice as fast as the Taylor series.

EXPERIMENTS WITH THE INVERSE OF A HILBERT MATRIX

Participating Member of Panel: S. E. Atta

Reference: "A Matrix Inversion and Linear-Equation-Solver Program," this report

The Hilbert matrix is well known for being "ill-conditioned" with respect to inversion. A Hilbert matrix of order n is defined by

$$Hn = \frac{1}{i+j-1} \ , \qquad i = 1, 2, \ldots, n \ , \quad j = 1, 2, \ldots, n \ .$$

[2] P. M. Hummel and C. L. Seebeck, Jr., *Amer. Math. Monthly* **56**, 243–247 (1949).

Available tables of exact inverses of these matrices are frequently used in the testing of approximating methods for inverting matrices on the Oracle. The test matrix used in these experiments was H_6.

A set of experiments was run to determine the extent to which propagated error in the original matrix affects its resultant inverse. Approximate inverses from different representations of the original elements of H_6, together with a measure of their accuracy, were computed. In each case the matrix inverted was not H_6, but H_6^* obtained from H_6 by rounding. A comparison of the results with the exact inverse of H_6 shows that a small error in the original matrix can have a large effect upon its inverse.

It was thought that the loss of significant digits during the process of triangularization could be avoided by making all rotational pivots of equal size. It was convenient to prescribe equal values to all rotational pivots by letting k take on other values. The pivots were set equal to the nth root of the determinant by first choosing k equal to +1 and computing the values and then the nth root of the determinant. The original matrix was again triangularized by rotating, with all pivots equal to the calculated nth root of the determinant. However, the results showed that nothing was gained in this experiment. Another test was made by setting all pivots equal to +1. The results of this experiment show that there was a small gain in accuracy.

The results of these experiments have been included in a paper written and submitted for publication in the *Journal of the Association for Computing Machinery.*

INVERSE OF A 190 × 190 MATRIX

Origin: C. J. McHargue, Metallurgy Division

Participating Members of Panel: M. R. Arnette, J. W. Givens, V. C. Klema

Status. – The inverse of the given matrix and four particular sets of solutions were calculated. Inspection of some of the results by the originator of the problem revealed the necessity for a more accurate approximation to the geometry of the experimental situations. This revision is understood to be under way, and further computations of the same type are anticipated.

CROSS SECTION FOR γ,n REACTION IN Be^9 AND IN C^{13}

Origin: E. Guth, G. de Saussure

Participating Member of Panel: M. E. Fulmer

Background and Status. – A calculation has been made of the cross sections for the γ,n reaction in Be^9 and in C^{13} expected on the basis of a model developed and designed by Guth and Mullin.[3]

[3] E. Guth and C. J. Mullin, *Phys. Rev.* **76**, 234 (1949).

THEORETICAL STUDIES IN NUMERICAL ANALYSIS. LOWER BOUND FOR SEPARATION OF CHARACTERISTIC VECTORS

Origin: J. W. Givens

Participating Member of Panel: J. W. Givens

The calculation of the characteristic vectors of a matrix by any iterative technique is made difficult by the fact that a vector and its transform may be very nearly equal, even when the characteristic values are readily distinguishable by the computational technique employed. Indeed, Wielandt has shown that, for x an arbitrary unit vector, y any vector not an exact multiple of x, and $\lambda_1, \lambda_2, \ldots, \lambda_n$ any n distinct complex numbers, there is a matrix A such that $Ax = y$ and the λ_i are the characteristic values of A.

In contrast, if the elements of the real matrix A are bounded, it has now been proved that the characteristic vectors belonging to distinct roots cannot be arbitrarily close: if $Ax = \lambda x$ for $\Sigma(x_i)^2 = 1$, $Az = \mu z$ for $z = x + y$ and x and y orthogonal, then

$$\max_i |y_i| \leqq |\mu - \lambda| \, (a + |\mu| n^{-1})^{-1} \, n^{-3/2} \ ,$$

where $\max_{i,j} |a_{ij}| = a$. An example shows that the exponent $-3/2$ in the lower bound cannot be replaced by an exponent greater than -1; hence the bound is reasonably realistic for matrices of small order.

VECTOR-BLOCK MATRIX CODES

Origin: J. W. Givens

Participating Members of Panel: M. R. Arnette, V. C. Klema

Background and Status. – An extended account of the complex of codes included under this heading which are in use or planned was given in the last semiannual report. Much of the work since then has gone into (1) input and output codes; (2) experimental error analysis done incidentally to testing the codes and rendering them error-free; and (3) rewriting and checking the codes for the new length of a block on magnetic tape. In some cases it has been desirable to write two versions of the codes, one for matrices with $\leqq 127$ rows and the other for those with >127 but $\leqq 255$ rows. Operating instructions for the main linear equations codes have been written but should not be released generally until the revised codes have been checked on the Oracle, which, in turn, requires rechecking subsidiary "service" routines.

A MATRIX INVERSION AND LINEAR-EQUATION-SOLVER PROGRAM

Origin: J. W. Givens

Participating Member of Panel: S. E. Atta

Reference: Mathematics Panel Semiannual Progress Report, ORNL-2037

Background and Status. – The method employed for inverting a matrix and for solving the associated simultaneous linear equations was devised by J. W. Givens.

This computational procedure is a direct one which triangularizes the matrix by applying successive plane rotations. Each rotation of the matrix A reduces an element below the diagonal to zero by multiplying the matrix A on the left by a matrix L_{ij}, which differs in only four places from the unit matrix, which is defined as

$$L_{ij} = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & C_{jj} & & & S_{ji} & \\ & & & & \ddots & & & \\ & & & & & 1 & & \\ & & & S_{ij} & & & C_{ii} & \\ & & & & & & & \ddots \\ & & & & & & & & 1 \end{pmatrix} ,$$

with

$$-S_{ij} = S_{ji} = S \ , \quad \text{and} \quad C_{ii} = C_{jj} = C \ ,$$

where i and j are the row and column indices, respectively, of the a_{ij} element which is to be reduced to zero. It is required that C and S be chosen such that

$$-Sa_{jj} + Ca_{ij} = 0 \ ,$$

$$C^2 + S^2 = k \ .$$

If k is chosen to equal +1, the value of the determinant is preserved. The triangularized matrix $A^{(n-1)}$ is equal to

$$\left(\prod_{j=1}^{n-1} \quad \prod_{i=1+j}^{n} \quad L_{ij} \right) A \ .$$

Final results are obtained by back-substitution. All the arithmetic operations are carried out in unpacked binary floating point. The entire matrix and the unit matrix are stored in the internal memory. The program is designed for the inverse of small-order matrices and the solution of simultaneous linear equations.

In summary, the program will accomplish the following:

1. either invert a matrix or solve systems of linear equations, or both;
2. determine, by checking the absolute size of the largest and smallest diagonal elements after triangularization, whether the matrix is singular or whether the linear equations are dependent;
3. check the inverse and the solutions for the approximate degree of accuracy;
4. take the input data from paper tape in packed or unpacked decimal floating point or take the input data from magnetic tape in unpacked binary floating point;
5. punch out the answers in converted packed or unpacked decimal floating point or the answers from magnetic tape in unpacked binary floating point;

6. designate the magnetic tapes to be used for the program, for the input data, and for the answers;
7. run on either single or double memory. (The largest matrix which can be inverted is a 21×21 on the double memory and a 14×14 on the single memory. The largest single system is for 29 equations on the double memory and for 20 on the single memory.)

TEST FOR THE EFFECT OF NEUTRON IRRADIATION ON LONGEVITY IN MICE

Origin: W. L. Russell, Biology Division

Participating Member of Panel: A. W. Kimball

Background and Status. – A small sample of F_1 mice from male parents exposed to neutrons from a nuclear detonation was presented for analysis of effects on longevity. This particular experiment was not designed for testing effects on longevity, and the data are highly nonorthogonal. Nevertheless, longevity data on progeny of irradiated mice are virtually nonexistent, and it was felt that an analysis of these data should be attempted.

After first testing for the existence of a cage effect and finding none, the data from different cages were pooled to yield the mean survival times shown in Table 1. The nonorthogonality is manifested by the variation in sample sizes from 1 to 53. Established methods for the analysis of two-way classifications with disproportionate subclass numbers yielded the analysis of variance shown in Table 2. It may be seen that the variance ratio for dose effect is 37,165/23,126 = 1.607, which is significant only at the 17.4% level. The other two effects obviously have variance ratios less than 1. It would appear that no dose effect can be established.

TABLE 1. MEAN SURVIVAL TIMES vs IRRADIATION DOSE

Dose (rep)	Survival (days)		
	Male	Female	Average
0	765.3 (53)*	821.9 (50)	792.3
22	773.1 (22)	739.8 (28)	754.5
50	688.2 (4)	744.0 (1)	699.4
83	746.3 (12)	693.9 (10)	722.5
96	663.0 (5)	729.0 (3)	687.8
131	787.0 (1)	724.0 (1)	755.5
Average	756.5	778.0	767.0

*Numbers in parentheses are sample sizes.

TABLE 2. ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Mean Square
Between sexes	1	17,261
Among doses	5	37,165
Sex-dose interaction	5	21,318
Within subclasses	178	23,126
Total	189	

This analysis, however, fails to make use of the quantitative information about dose. Accordingly, a straight line was fitted to the three dose points having the largest number of mice, that is, doses of 0, 22, and 83 rep. This analysis yielded the following results:

Intercept	787.4
Slope	−0.8808
Accident variance	13,094
Standard error of slope	0.4270

The fact that the residual variance is actually less than that with the subclass mean square in Table 2 indicates that there is no reason to suspect nonlinearity in the dose range from 0 to 83 rep. The larger mean square was used to compute the variance of the slope, and a two-sided t-test showed that the slope differs significantly from zero at the 4% level. Since a one-sided test would be more appropriate, the significance level should probably be taken as 2%.

Thus there is good evidence of an effect of neutron irradiation on longevity in the F_1 generation. A bold interpretation might be that the reduction in life span is of the order of 0.9 day/rep, although the relatively large standard error of the estimated slope should temper such conjectures.

A MODEL FOR THE MULTIPLE-CAUSE-OF-DEATH PROBLEM

Origin: A. C. Upton, Biology Division

Participating Members of Panel: A. W. Kimball, G. J. Atta

Background. – The general problem is illustrated most easily by reference to a particular experiment. A group of mice, containing equal numbers of both sexes, is exposed to a sublethal dose of x radiation. The group is followed until all mice have died, and then each animal is classified as having died as a result of thymic leukemia,

myeloid leukemia, or other causes. The biologist wishes to test whether the incidence of myeloid leukemia is the same for both sexes. There are, however, two complicating factors. Thymic leukemia generally occurs earlier in life than does myeloid leukemia and is much more prevalent in one sex than in the other. Thus in both sexes there are animals which died of thymic leukemia, before they could manifest myeloid leukemia, but this number is not the same for both sex groups. The problem is to estimate what the incidences of myeloid leukemia would be if thymic leukemia were eliminated as a cause of death.

Status. – In general let there be r different causes of death. If the only cause of death operating were cause α ($\alpha = 1, \ldots, r$), it can be assumed that the probability density for a death due to this cause at time t is

$$(1) \qquad \begin{aligned} p_\alpha^\star(t) &= 2k_\alpha(t - \mu_\alpha)\, e^{-k_\alpha(t-\mu_\alpha)^2}, & \mu_\alpha < t < \infty, \\ &= 0, & 0 < t \leqq \mu_\alpha. \end{aligned}$$

Then the probability that an animal has not died of cause α up to time t is

$$\int_t^\infty p_\alpha^\star(t)\, dt = e^{-k_\alpha(t-\mu_\alpha)^2}.$$

Now, if all causes of death are operating independently, the probability density, $p_\alpha(t)$, for a death due to cause α at time t is $p_\alpha^\star(t)$ multiplied by the probability that the animal has not died of any other cause up to time t; that is,

$$(2) \qquad \begin{aligned} p_\alpha(t) &= 2k_\alpha(t - \mu_\alpha)\, e^{-k_\alpha(t-\mu_\alpha)^2} \prod_{i \neq \alpha} \left[e^{-k_i(t-\mu_i)^2} \right] \\ &= 2k_\alpha(t - \mu_\alpha)\, e^{-\sum_{i=1}^{r} k_i(t-\mu_i)^2}. \end{aligned}$$

Furthermore, the incidence of death due to cause $\alpha (I_\alpha)$ is given by

$$(3) \qquad I_\alpha = \int_0^\infty p_\alpha(t)\, dt,$$

where caution must be taken to observe the ranges of integration which are specified in Eq. 1. If the parameters k and μ were known for every cause of death, it would, in principle, be possible to compute the incidence for any cause of death operating simultaneously with any subset of the original r causes of death. In particular, this could be done for the leukemia problem.

Consider a sample of size N in which n_α animals die at times $t_{\alpha j}$ ($j = 1, \ldots, n_\alpha$) from cause α, where

$$N = \sum_{\alpha=1}^{r} n_\alpha.$$

On the basis of Eq. 2, the probability of such a sample is

$$P(S) = 2^N k_1^{n_1} \ldots k_r^{n_r} \prod_{\alpha=1}^{r} \prod_{j=1}^{n_\alpha} (t_{\alpha j} - \mu_\alpha)\, e^{-\sum_{\alpha=1}^{r} \sum_{j=1}^{n_\alpha} \sum_{i=1}^{r} k_i (t_{\alpha j} - \mu_i)^2} .$$

The maximum likelihood equations for the $2r$ parameters to be estimated from this sample can be divided into r pairs of equations, such that each pair contains two, and only two, parameters and that the parameters occurring in any one pair occur in no other pair. One such pair, corresponding, say, to the pth cause of death is

$$\frac{n_p}{k_p} - \sum_{\alpha=1}^{r} \sum_{j=1}^{n_\alpha} (t_{\alpha j} - \mu_p)^2 = 0 ,$$

$$-\sum_{j=1}^{n_p} \left(\frac{1}{t_{pj} - \mu_p} \right) + 2k_p \sum_{\alpha=1}^{r} \sum_{j=1}^{n_\alpha} (t_{\alpha j} - \mu_p) = 0 .$$

The first equation can be solved for k_p and substituted in the second equation, which can then be solved for μ_p by trial and error.

This method of estimation was tried on several sets of leukemia data, and good fits to the survival curves were obtained in most cases. Unfortunately, the matrix

$$-\begin{pmatrix} \mathcal{E}\left[\dfrac{\partial^2 \log P(S)}{\partial k_p^2}\right] & \mathcal{E}\left[\dfrac{\partial^2 \log P(S)}{\partial k_p\, \partial \mu_p}\right] \\ \mathcal{E}\left[\dfrac{\partial^2 \log P(S)}{\partial k_p\, \partial \mu_p}\right] & \mathcal{E}\left[\dfrac{\partial^2 \log P(S)}{\partial \mu_p^2}\right] \end{pmatrix} ,$$

the inverse of which is the asymptotic variance-covariance matrix for the estimates of k_p and μ_p, appears to have an infinite element in the lower right corner. This implies that the estimate of μ_p is a hyperefficient statistic; therefore conventional methods for obtaining asymptotic variances are useless. Incidences based on Eq. 3 will not be computed until some method is found for obtaining error estimates. As an alternative, attempts are being made to modify certain nonparametric methods, with the hope that they can be applied to this problem. This latter approach, if successful, would be preferable, since it would obviate the need for assumptions such as those embodied in Eq. 1.

A VARIANCE-STABILIZING TRANSFORMATION FOR CERTAIN *PARAMECIUM* EXPERIMENTS

Origin: R. F. Kimball, Biology Division

Participating Members of Panel: A. W. Kimball, G. J. Atta

Background and Status. – Certain *Paramecium* experiments are based on proportions of the form a/n, where n is the total number of animals and a is the number of normal animals ($a \leqq n$). It has been shown that if a/n has an expected value of θ, the variance of a/n is given by

$$V\left(\frac{a}{n}\right) = \frac{\theta}{n}\left(1 - \frac{n+1}{2}\theta + \frac{n-1}{2}\sqrt{\theta}\right).$$

Proportions of this type, in which the variance depends on the mean, are not amenable to treatment by analysis of variance unless some way is found to remove the dependency of $V(a/n)$ on θ.

One theorem is that if an estimate T of a parameter θ has a variance of the form $g(\theta)/n$, if certain continuity conditions are satisfied, and if

$$f(\theta) = \int \frac{d\theta}{\sqrt{g(\theta)}}$$

exists, then $f(T)$ will have a variance approximately equal to $1/n$. In the present problem,

$$f(\theta) = \int \frac{d\theta}{\sqrt{\theta\left(1 - \frac{n+1}{2}\theta + \frac{n-1}{2}\sqrt{\theta}\right)}}$$

$$= \frac{2\sqrt{2}}{n+1}\sin^{-1}\frac{2(n+1)x - (n-1)}{n+3} + \text{const},$$

where $x = \sqrt{a/n}$. It is perhaps convenient to take the constant equal to

$$\frac{2\sqrt{2}}{\sqrt{n+1}}\sin^{-1}\frac{n-1}{n+1},$$

since then $f(0) = 0$, and $f(x) > 0$ for $x > 0$. In practice it will be simple to use the transformation

$$f^*(x) = \frac{\sqrt{n+1}}{2\sqrt{2}}f(x)$$

$$= \sin^{-1}\left(\frac{n-1}{n+3}\right) + \sin^{-1}\frac{2(n+1)x - (n-1)}{n+3};$$

whereupon

$$V[f^*(x)] \sim \frac{n+1}{8n} .$$

This transformation was applied to several sets of data in which $n = 25$. The coefficient of variation among the sample variances was reduced from 60 to 14%, and the average transformed variance was 0.188, slightly higher than its expected value of 26/200 = 0.13. The transformation is apparently successful in stabilizing the variance.

CURVE-PLOTTER CODE EDIT

Participating Member of Panel: C. L. Gerberich

Status. – The curve-plotter code edit has been applied to the manufacture of readable, reliable copies of production codes. This code loads two tapes, an input tape and the actual program code, and combines the two into a readable code that is photographed on the curve plotter. The first information that the input tape supplies is a list of up to 32 characters that will appear on the first line of every photograph taken by the curve plotter. This line is used for identification of the film and, therefore, usually includes such items as the programmer's name, the problem number, and the date. The last nine characters in the first line of every photograph are reserved for the page number. The input tape supplies the page number for the first photograph, and this number is advanced by 1 on each of the following photographs. The next information on the input tape is n lines of, at most, 41 characters, each of which will be used as a heading. The number of lines, n, is not limited. The code will write the heading information at the rate of 31 lines per photograph until all n lines are written. The heading information usually includes such items as the title of the problem, a short description, and the operational instructions. After the heading material is complete, the input tape supplies the address of the first word on the program tape. This address and all the following addresses form a column on the left of the photograph. The next column is made up of the code words on the program tape. The code also makes provision for three other columns. The first and second are for labeling the beginnings of the boxes on the accompanying flow charts. The first is for the left-hand order, the second for the right-hand order. Each of these columns contains three characters. The third column is for general remarks and contains up to 15 characters. The input tape lists addresses at which there is extra material, which of the three extra columns are to be used, and what the extra characters are.

At the present time the code uses the key words for the curve plotter as listed in *Oracle Memo 4*. Therefore, any character on that list can be used, and each is given a hex address from 00 to 4D. Thus the code can refer to any character on the list by

using the two-character hex address associated with it. The addresses $4E$ to FF can be used for future expansion or for private symbols that may be added by the individual programmer. The program has been used for several codes and was found to be beneficial. The experience gained in using it on these codes has suggested several modifications which will increase the ease of operation and its usefulness.

DAMAGE TO THIN SLABS OF TISSUE FROM NEUTRON IRRADIATION

Origin: W. S. Snyder, Health Physics Division

Participating Member of Panel: N. M. Dismuke

Reference: Mathematics Panel Semiannual Progress Report, ORNL-2037

Background and Status. – In order to determine the slab thickness beyond which the dose is equivalent to a first-collision dose, a thin-slab dose study has been undertaken. Neutron and gamma-ray histories are being computed by using the codes of earlier reported work (see reference, p 13). Histories for 10-, 5-, 2.5-, 1-, and 0.5-cm slabs have been calculated. Analysis of these results will determine whether data for thinner slabs will be necessary.

DETERMINATION OF THE BREIT-WIGNER CONSTANTS FOR XENON

Origin: E. C. Smith

Participating Member of Panel: N. M. Dismuke

Background. – In order to compare the Breit-Wigner constants for the xenon total neutron cross section determined with and without the Doppler correction, a general fitting code was written for the Oracle. The code applies the Garwood fitting procedure.

Status. – The computation with Doppler correction has been completed. There remains the computation without Doppler correction and the determination of the average values for reactor neutrons of total and capture xenon cross sections.

FISSION-PRODUCT PROBLEM

Origin: R. A. Charpie, Director's Division

Participating Member of Panel: M. F. Todd

Background and Status. – The atomic concentrations of the U^{235} fission products have been computed as functions of the parameters thermal-neutron flux, irradiation time, and cooling time. These numbers have been stored on magnetic tape, and the normalized values of the concentrations were put out in digital form on the curve plotter. Approximately 2500 pictures, 200,000 numbers, have been given to J. O. Blomeke for editing and publication.

In addition to the values for individual fission products, summations of the atomic concentration, activity, gamma power, total power, poisoning, and activity by gamma energy ranges for the entire fission-product spectrum have been completed.

HAND COMPUTATIONS

Origin: J. Neufeld, Health Physics Division

Participating Members of Panel: G. L. Broyles, C. D. Griffies, F. R. Mance

Calculations were made of W in the equation for the average charge of heavy particles,

$$W = \left[\frac{21.92 \times A(Z_1^*)^{5/3}\, Z_1^{4/3}\, Z_2^{1/3}\, N}{(Z_1^* + 1)^4 - 2\pi Z_2^{1/3}\, NAZ_1^*\, Z_1^{4/3}}\right]^{3/2},$$

for various values of Z_1^*A, Z_1, Z_2, and N.

CALCULATION OF STRESSES IN PIPING SYSTEM

Origin: M. I. Lundin, Reactor Experimental Engineering Division

Participating Member of Panel: E. C. Long

References: Mathematics Panel Semiannual Progress Reports, ORNL-1928, -2037

Background and Status. – A program has been completed for computing stresses in a four-branch line which recombines into one line with an intermediate restraint. This involves solving a system of 48 equations and applying the solution (forces and moments) to compute the stresses.

CYLINDER ACTIVATION AND SOLID-ANGLE CALCULATIONS

Origin: R. L. Macklin, A. V. H. Masket, and H. W. Schmitt, Physics Division

Participating Member of Panel: M. F. Todd

Reference: Mathematics Panel Semiannual Progress Report, ORNL-2037

Background and Status. – As an outgrowth of the Shell Transmission Experiment it was deemed worth while to prepare an extensive compilation of cylinder-activation and solid-angle data.

The following values of the parameters were used:

$0 \leqq r \leqq 6$	At intervals of 0.1
$6.5 \leqq r \leqq 16$	At intervals of 0.5
$17 \leqq r \leqq 36$	At intervals of 1
$0.1 \leqq h \leqq 5$	At intervals of 0.1
$5.2 \leqq h \leqq 10$	At intervals of 0.2
$11 \leqq h \leqq 20$	At intervals of 1
$22 \leqq h \leqq 100$	At intervals of 2

where h is the height of a cylinder of unit radius and r is the distance from its axis.

In order to correct for the significance loss which occurred for certain ranges of the parameters, it was necessary to evaluate the activation by a double integral and to reformulate the single integral for the solid angle. The double integral used is

$$I = 2\pi \int_{z=0}^{h} \int_{\rho=0}^{1} \frac{\rho \, d\rho \, dz}{\sqrt{(z^2 + \rho^2 + r^2)^2 - 4\rho^2 r^2}} .$$

An alternative formulation of the activation integral which was found to give more consistent results for the smallest values of h and for $0 \leqq r \leqq 1.3$ is

$$I = \int_0^{\pi} \left(h \ln \frac{c^2 + h^2}{d^2 + h^2} + 2c \tan^{-1} \frac{h}{c} - 2d \tan^{-1} \frac{h}{d} \right) d\phi ,$$

where

$$c = r \cos \phi + \sqrt{1 - r^2 \sin^2 \phi} , \qquad d = 0 , \qquad \text{for } r < 1 ;$$

$$I = \int_0^{\tan^{-1}[1/(r^2-1)]} \left(h \ln \frac{c^2 + h^2}{d^2 + h^2} + 2c \tan^{-1} \frac{h}{c} - 2d \tan^{-1} \frac{h}{d} \right) d\phi ,$$

where

$$c = r \cos \phi + \sqrt{1 - r^2 \sin^2 \phi} , \qquad d = r \cos \phi - \sqrt{1 - r^2 \sin^2 \phi} , \qquad \text{for } r \geqq 1 .$$

The solid-angle integrals for the circular disk become:

$$\Omega(1,r,h) = 2 \int_0^{\pi} \left(\frac{f_1^2}{1 + f_1^2 + \sqrt{1 + f_1^2}} \right) d\phi ,$$

where

$$f_1 = \frac{r}{h} \left[\cos \phi + \sqrt{\left(\frac{1}{r}\right)^2 - \sin^2 \phi} \right] , \qquad \text{for } r < 1 ;$$

$$\Omega(1,r,h) = 2 \int_0^{\pi/2} \left(\frac{f^2}{1 + f^2 + \sqrt{1 + f^2}} \right) d\phi ,$$

where

$$f = \frac{2 \cos \phi}{h} , \qquad \text{for } r = 1 ;$$

$$\Omega(1,r,h) = 8\left(\frac{r}{h}\right)^2 \int_0^{\sin^{-1}(1/r)} \frac{\cos \phi \sqrt{(1/r)^2 - \sin^2 \phi}}{\sqrt{(1 + f_1^2)(1 + f_2^2)}\left(\sqrt{1 + f_1^2} + \sqrt{1 + f_2^2}\right)} d\phi ,$$

where

$$f_1 = \frac{r}{b}\left[\cos\phi + \sqrt{\left(\frac{1}{r}\right)^2 - \sin^2\phi}\right] ,$$

$$f_2 = \frac{r}{b}\left[\cos\phi - \sqrt{\left(\frac{1}{r}\right)^2 - \sin^2\phi}\right] ,$$

for $r > 1$.

The solid-angle integral for the lateral surface of a right circular cylinder becomes

$$\Omega(1,b,r) = 2\int_0^{\sin^{-1}(1/r)} \left(1 + f_2^2\right)^{-1/2} d\phi ,$$

where

$$f_2 = \frac{r}{b}\left[\cos\phi - \sqrt{\left(\frac{1}{r}\right)^2 - \sin^2\phi}\right] .$$

ECHELLOGRAM DATA REDUCTION

Origin: K. L. Vander Sluis and P. M. Griffin, Stable Isotopes Division

Participating Member of Panel: J. H. Vander Sluis

Reference: Mathematics Panel Semiannual Progress Report, ORNL-2037

Background. – The previously reported routine (see reference, p 11–12) now in use on production problems carries out the reduction of Zeeman information for individual transitions. For each transition this routine calculates the vacuum wave-number of the transition (the energy difference of the levels involved), computes the Landé g factor and its probable error, and lists the quantum number for each of the two levels involved. With this information in the Oracle, it is possible to proceed to the basic problem of establishing the energy-level array which gives rise to the observed spectrum. The proposed procedure is the following:

1. sorting by vacuum wave-numbers, which will provide a means to average all information pertaining to individual transitions;
2. sorting according to the quantum number J and the Landé g factor, which will define an energy level (J,g) and the relative position of this level with respect to a chain of other levels $(\bar{\sigma}_i)$;
3. forming all sums and differences of $(\bar{\sigma}_i)$ in each chain and intercomparing, which will allow the interlocking of chains into an array of energy levels;
4. making use of selection rules and the energy-level array, which will permit the wave-number list to be extended and compared.

This second phase of the problem necessitated consideration of various methods of sorting data in order to arrive at a sorting subroutine which can be placed in the subroutine library. The procedure decided upon was one described by Goldstine and von Neumann.[4] A modification of this method, used earlier by R. R. Coveyou and R. R. Bate in a less general code, will sort one block of data in the fast memory, using two blocks for intermediate and final storage. Since many blocks of data may be sorted, von Neumann's technique for the fast memory was extended for magnetic tape, using one tape for beginning and final storage and two tapes for intermediate storage. The code in more general subroutines will allow identification words to be carried along with the numbers to be sorted; will permit the possibility of sorting as many as 60,000 numbers or, with a minor change in the code, any amount desired, subject only to the storage capacity of the tapes; will allow incomplete blocks to be sorted; and will permit the programmer to choose the tapes to be used.

Status. – The sorting subroutines are nearly completed, and work is progressing on the specific problem.

VARIANCE CALCULATIONS IN A *DROSOPHILA* EXPERIMENT

Origin: D. L. Lindsley, Biology Division

Participating Members of Panel: A. W. Kimball, G. J. Atta

Background and Status. – In a *Drosophila* experiment involving losses in two types of chromosomes, the data may be represented in four equations:

$$x = p_1 , \tag{1}$$

$$y = p_2 , \tag{2}$$

$$\frac{1 - x}{(1 - x) + x(1 - a)} = p_3 , \tag{3}$$

$$\frac{y(1 - a)}{y(1 - a) + (1 - y)(1 - \alpha)} = p_4 , \tag{4}$$

where x, y, a, and α are four parameters to be estimated and where the p_i are binomial variates based on n_i observations and having expectations given by the left sides of the equations. Estimates of x and y are given directly by Eqs. 1 and 2 and have the usual binomial variances p_1q_1/n_1 and p_2q_2/n_2, where $q_i = 1 - p_i$. Substitution of Eqs. 1 and 2 into Eqs. 3 and 4 yields the following estimates for a and α:

[4]H. H. Goldstine and J. von Neumann, *Planning and Coding Problems for an Electronic Computing Instrument*, part II, vol II, p 49–66, Institute for Advanced Study, Princeton, 1948.

$$\hat{a} = \frac{p_1 - q_3}{p_1 p_3} ,$$

$$\hat{\alpha} = \frac{(p_4 - p_2) + \hat{a} p_2 q_4}{p_4 q_2} .$$

Computation of variances for $\hat{a}$ and $\hat{\alpha}$ is more difficult. It can be accomplished, however, by successive application of the following variance formulas, where u and v are random variables not necessarily independent and have variances σ_u^2 and σ_v^2, covariance σ_{uv}, and expected values $\bar{u}$ and $\bar{v}$:

$$V(u + v) = \sigma_u^2 + \sigma_v^2 \pm 2\sigma_{uv} ,$$

$$V(uv) \sim \bar{u}^2 \sigma_v^2 + \bar{v}^2 \sigma_u^2 + 2\bar{u}\bar{v}\sigma_{uv} ,$$

$$V\left(\frac{u}{v}\right) \sim \left(\frac{\bar{u}}{\bar{v}}\right)^2 \left(\frac{\sigma_u^2}{\bar{u}^2} + \frac{\sigma_v^2}{\bar{v}^2} - \frac{2\sigma_{uv}}{\bar{u}\bar{v}}\right) .$$

When this procedure is followed, the variances of a and α are found to be

$$V(\hat{a}) \sim \frac{q_1 q_3}{(p_1 p_2)^2} \left(\frac{q_3}{n_1 p_1} + \frac{q_1}{n_3 p_3}\right) ,$$

$$V(\hat{\alpha}) \sim \frac{1}{(p_4 q_2)^2} \left[\frac{p_4 q_4}{n_4} (1 - \hat{a} p_2 - \hat{\alpha} q_2)^2 + \frac{p_2 q_2}{n_2} (1 - \hat{a} q_4 - \hat{\alpha} p_4)^2 + (p_2 q_4)^2 V(\hat{a})\right] ,$$

where expected values have been replaced by their sample estimates.