UNCLASSIFIED

OAK RIDGE NATIONAL LABORATORY LIBRARY
3 4456 0022223 9

ORNL-2283
Progress
cy. 6

CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION

MATHEMATICS PANEL

SEMIANNUAL PROGRESS REPORT

FOR PERIOD ENDING FEBRUARY 28, 1957

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 X · OAK RIDGE, TENNESSEE

UNCLASSIFIED

Printed in USA. Price 30 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-2283

Contract No. W-7405-eng-26

MATHEMATICS PANEL

SEMIANNUAL PROGRESS REPORT

for

Period Ending February 28, 1957

A. S. Householder, Chief

DATE ISSUED

JUN 3 1957

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

UNCLASSIFIED

OAK RIDGE NATIONAL LABORATORY LIBRARY

3 4456 0022223 9

ORNL-2283
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. J. Harrison
43. N. M. Dismuke
44. A. H. Culkowski
45. A. W. Kimball
46. M. R. Arnette
47. D. D. Cowen
48. P. M. Reyling
49. A. C. Downing, Jr.
50. J. H. Vander Sluis
51. G. J. Atta
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. M. A. Kastenbaum
61. J. P. Kelly (K-25)
62. S. E. Atta
63. G. W. Medlin
64. F. J. Witt
65. D. A. Gardiner
66. ORNL – Y-12 Technical Library, Document Reference Section

EXTERNAL DISTRIBUTION

67. A. E. S. Green, Florida State University
68. R. F. Bacher, California Institute of Technology
69. Division of Research and Development, AEC, ORO
70–644. Given distribution as shown in TID-4500 (13th ed.) under Physics and Mathematics category (100 copies – OTS)

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
ORNL-2134	Period Ending June 30, 1956

CONTENTS

SUMMARY 1

NUMERICAL ANALYSIS RESEARCH 3

- Fredholm Integral Equations 3
- Quotient-Difference Algorithm 3
- Calculation of the Characteristic Roots and Vectors of Matrices 5
- Calculation of the Matrix Product $A \cdot B^T$ 6
- Calculation of the Transpose of a Matrix 6
- Characteristic Values and Vectors of Real Symmetric Matrices 6
- Matrix Inversion and Linear Equation Solver 7
- Sorting Routines 7
- Vector-Block Matrix Codes 9
- Additions to the Subroutine Library 9

STATISTICS 11

- Disease-Incidence Estimation in Populations Subject to Multiple Causes of Death 11
- Inverse of a Certain Patterned Matrix 13
- A Problem in Estimation by the Method of Maximum Likelihood. I 16
- A Problem in Estimation by the Method of Maximum Likelihood. II 19
- A Nonlinear Dose-Effect Fitting Problem 20
- Estimation of the Diameter of the *Habrobracon* Nucleus 21

PROBLEMS 22

- Application of Standard Codes 22
- Short-lived Fission-Product Gamma Rays 23
- Echellogram Data Reduction 26
- Determination of the Breit-Wigner Constants for Xenon 26
- Characteristics of Uranyl Sulfate Solutions 26

ORACLE OPERATIONS, SERVICE ROUTINES, AND EDUCATION 27

- Operation of the Oracle 27
- Curve-Plotter Dump 27
- Curve-Plotter Post-Mortem 28
- Tracer Routine 29
- Oracle Programing Class 30

PUBLICATIONS 31

MATHEMATICS PANEL SEMIANNUAL PROGRESS REPORT

SUMMARY

The demands of the Laboratory staff for high-speed computation rose in the last eight months to a level requiring the operation of the Oracle on a 24-hr-day, 7-day-week basis. Since further expansion in hours was impossible, attention had to be given to more efficient use of the available computing time. To this end, service routines were prepared in order to make the photographic output easier to use.

Attention was also given to rewriting some of the basic routines which were already out of date because of the many alterations and additions which have been made to the Oracle since its original installation. One of these is the sequence of routines originally conceived by J. W. Givens for obtaining the characteristic values and vectors of a symmetric matrix. The principal interest in such computations lies with those scientists investigating molecular structure. The problems they wish to have solved, however, often require the characteristic values and vectors of the product of two symmetric matrices or involve inverse characteristic-value problems such as those which arise when the force constants of a molecule are to be predicted. An extensive effort was undertaken to prepare efficient routines for handling such problems, and, at the same time, an investigation was undertaken to discover a suitable numerical technique for overcoming the difficulties associated with the computation of a complete set of characteristic vectors when multiple characteristic values occur. Experimentation with the finished routines indicates that further study and programing will be necessary before the latter problem is resolved.

In the past the principal effort toward making the Oracle more useful has been devoted to the preparation of subroutines for evaluating functions and functionals and to the preparation of interpretive codes for performing arithmetic operations not directly available in the Oracle. This work can never be called complete, but, nevertheless, the progress in this direction has been so great that it appears desirable to change the emphasis so as to devote the greater portion of the available effort to the preparation of routines for solving complete mathematical problems. Such routines – or "canned codes," as they may be called – have only been written in the past when a particular need arose, with the unfortunate result that they were not prepared in the greatest generality and, hence, were of limited usefulness. Aside from the characteristic-value routines, the first of these canned codes will be a general-purpose linear least-square fitting routine. The programing is about 80% complete.

The theoretical research in the numerical analysis of truncation error and computational stability in the solution of partial differential equations has been continued during this period. In addition, a method for solving certain Fredholm integral equations

by numerically computing upper and lower step functions which serve as bounds for the solution has been developed and studied.

Some experimental programing has been undertaken to test the practicality of using the quotient-difference algorithm for obtaining the roots of high-order polynomials. The results are very encouraging. This work leads naturally into the study of methods for obtaining characteristic values of nonsymmetric matrices. Another approach to this latter problem which has been investigated is the power method.

There has been a noticeable increase in requests for statistical assistance from the Biology and the ARE Divisions of the Laboratory. Principal research efforts have been directed toward problems in the design of experiments, the solution of certain maximum-likelihood problems in radiation genetic experiments, finding internal estimates of specific-locus mutation rates, and the estimation of disease incidences in popluations subject to multiple causes of death. Portions of this work which have been completed are reported in detail. Computations for the Greenhouse experiment are nearing completion.

The utilization of the Oracle during the eight-month period beginning July 1, 1956, and ending February 28, 1957, may be summarized as follows: 75% of the available computing time was devoted to routines programed by persons not belonging to the Mathematics Panel. The Mathematics Panel used 12% of the time for mathematical research and the preparation of subroutines and canned codes, and 13% for solving problems originating outside the Mathematics Panel.

NUMERICAL ANALYSIS RESEARCH

FREDHOLM INTEGRAL EQUATIONS[1]

Origin: W. S. Snyder, Health Physics Division

Participating Member of Panel: C. L. Gerberich

Background and Status. – A technique for the solution of some types of Fredholm integral equations was suggested by W. S. Snyder. The object of this method is to find upper and lower bounds for the solution of

$$f(x) = s(x) + \lambda \int_a^b K(x,y)\, f(y)\, dy \ .$$

The technique, which is referred to as the method of upper and lower functions, develops methods of producing upper and lower bounds which are easily recognized as such. A function $g(x)$ is called an "upper function" if for a nonnegative kernel

$$g(x) \geqq g_1(x) \ , \qquad \text{for all } x \in [a,b] \ , \tag{1}$$

where

$$g_1(x) = s(x) + \lambda \int_a^b K(x,y)\, g(y)\, dy \ . \tag{2}$$

A function $h(x)$ is called a "lower function" if for a nonnegative kernel

$$h(x) \leqq h_1(x) \ , \qquad \text{for all } x \in [a,b] \ , \tag{3}$$

where

$$h_1(x) = s(x) + \lambda \int_a^b K(x,y)\, h(y)\, dy \ . \tag{4}$$

It can be shown that with certain conditions on the kernel the upper (lower) functions are upper (lower) bounds for the solution. As in many numerical methods, the original equation is replaced by a system of linear equations; the technique of upper and lower functions may be extended to such a system.

Some numerical examples have been calculated on the Oracle. At the present time, changes are being made in the original programs so that they can be used for a problem on the dose in various-sized spheres.

QUOTIENT-DIFFERENCE ALGORITHM

Participating Member of Panel: A. A. Grau

Background. – The quotient-difference algorithm (Q.D. algorithm) was developed by Rutishauser;[2] it may be considered as an extension of a result obtained by D. Bernoulli

---

[1] C. L. Gerberich, *On the Solution of a Fredholm Integral Equation*, ORNL-2241 (Dec. 19, 1956).

[2] H. Rutishauser, *Z. angew. Math. u. Phys.* 5, 233–251 (1954).

which locates the pole of smallest modulus of a meromorphic function. Among its possible applications are the solution of a polynomial equation and the determination of the eigenvalues of a (nonsymmetric) matrix in triple diagonal form.

Given a polynomial equation

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_n = 0 , \qquad a_i \neq 0 \; (i = 0, 1, \ldots, n) ,$$

quantities $q_\nu^{(i)}$, $e_\nu^{(i)}$ may be uniquely determined, provided they exist, from the relations

$$q_1^{(-1)} = -\frac{a_1}{a_0} , \qquad q_2^{(-2)} = q_3^{(-3)} = \ldots = q_n^{(-n)} = 0 ,$$

(1)

$$e_i^{(-i)} = +\frac{a_{i+1}}{a_i} ; \qquad e_0^{(k)} = e_n^{(k)} = 0 \qquad \text{(for all } k) ,$$

and the quotient-difference recursion relations (Q.D. algorithm)

$$q_\nu^{(i+1)} = q_\nu^{(i)} + e_\nu^{(i)} - e_{\nu+1}^{(i)} ,$$

(2)

$$e_\nu^{(i+1)} = \frac{e_\nu^{(i)} \, q_{\nu+1}^{(i)}}{q_\nu^{(i+1)}} .$$

In the case of a matrix of the type

$$\begin{pmatrix} \alpha_1 & \beta_1 & 0 & \ldots & 0 & 0 \\ 1 & \alpha_2 & \beta_2 & \ldots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots & \alpha_{n-1} & \beta_{n-1} \\ 0 & 0 & 0 & \ldots & 1 & \alpha_n \end{pmatrix} ,$$

a similar double sequence may be determined by setting

$$e_0^{(i)} = 0 \qquad \text{(for all } i) ,$$

$$q_i^{(0)} = \alpha_i - e_{i-1}^{(0)} ,$$

$$e_i^{(0)} = \frac{\beta_i}{q_i^{(0)}} ,$$

and by using Eqs. 2.

Suppose $\lambda_1, \ldots, \lambda_n$ represent the roots of the equation or eigenvalues of the matrix, ordered so that

$$|\lambda_1| \geqq |\lambda_2| \geqq \ldots \geqq |\lambda_n| .$$

The utility of the algorithm arises from the following theorems, on which the present work is based:

1. If $|\lambda_\nu| \neq |\lambda_{\nu+1}|$, $\lim_{i\to\infty} e_\nu^{(i)} = 0$.

2. If $|\lambda_{\nu-1}| \neq |\lambda_\nu| \neq |\lambda_{\nu+1}|$, $\lim_{i\to\infty} q_\nu^{(i)} = \lambda_\nu$.

3. If $|\lambda_{\nu-1}| \neq |\lambda_\nu| = |\lambda_{\nu+1}| \neq |\lambda_{\nu+2}|$,

$$\lim_{i\to\infty} \left[q_\nu^{(i)} + e_\nu^{(i)} + q_{\nu+1}^{(i)}\right] = \lambda_\nu + \lambda_{\nu+1} ,$$

and

$$\lim_{i\to\infty} \left[q_\nu^{(i)} q_{\nu+1}^{(i)}\right] = \lambda_\nu \lambda_{\nu+1} .$$

In cases where the algorithm fails because at some point division by zero is required, a translation of coordinates

$$x = x' - h$$

may be performed on the equation or matrix, the algorithm carried out, and the inverse translation effected to find the roots.

Status. – Experimental codes using the above principles have been set up for use on the Oracle. These employ decimal input and output, fixed-point single-precision operation with decimal scaling on overflow, and are set up to find both real and imaginary roots of equations with real coefficients or matrices with real elements. If, after 256 lines of algorithm, the roots are not yet separated, the last line of the algorithm may be punched out for inspection. The capacity of the fast memory would allow the handling of a matrix or equation of order 300 by this method. The codes also include the option of an accelerated form of the algorithm, which is useful when all roots are real and distinct.

The polynomial code has been used to solve sixth- and seventh-degree equations which had complex roots and which originated at the Laboratory.

The matrix code has been tested to a limited degree on matrices of orders including the twenty-third. Further intensive testing is planned.

CALCULATION OF THE CHARACTERISTIC ROOTS AND VECTORS OF MATRICES

Origin: A. S. Householder

Participating Member of Panel: G. W. Medlin

Background. – The power method[3] may be used in the calculation of the characteristic roots and vectors of matrices. The speed of convergence depends upon the ratio

$$\frac{|\lambda_1|}{|\lambda_2|} \qquad (|\lambda_1| > |\lambda_2| \geqq |\lambda_\mu| , \quad \mu \neq 1, 2)$$

of the moduli of the two largest characteristic values. This speed may be improved by

[3] A. S. Householder, *Principles of Numerical Analysis*, p 162–166, McGraw-Hill, New York, 1953.

a translation. Special devices are available for complex roots and for roots of equal modulus.

Status. – A program is now available for getting the root of greatest modulus and its vector. Two methods for getting the other roots are being evaluated.

CALCULATION OF THE MATRIX PRODUCT $A \cdot B^T$

Origin: Mathematics Panel

Participating Member of Panel: G. W. Medlin

Background. – Let A and B be arbitrary real matrices. Denote by A_r, A_c, B_r, B_c the number of rows and columns of A and B, respectively, and by B^T the transpose of B. If $A_c = B_c$, then the multiplication $A \cdot B^T$ is defined.

Status. – Let z denote the number of zeros needed to fill out the last block of magnetic tape in writing A_c words, that is, $A_c + z = k \cdot 128$ $(0 \leqq z < 128)$. A program has been completed for the multiplication $A \cdot B^T$ when $2A_c + z + B_r < 1750$.

CALCULATION OF THE TRANSPOSE OF A MATRIX

Origin: Mathematics Panel

Participating Member of Panel: G. W. Medlin

Background. – In computing the matrix product $A \cdot B^T$, where B^T is the transpose of B, it is advantageous to have the matrix A stored on magnetic tape by rows, and B by columns.

Status. – A program has been completed for computing the transpose of a matrix.

CHARACTERISTIC VALUES AND VECTORS OF REAL SYMMETRIC MATRICES

Participating Members of Panel: N. M. Dismuke, A. C. Downing, T. W. Hildebrandt,[4] E. C. Long, B. J. Osborne

Background and Status. – A series of programs for the Oracle has been written for the automatic solution of several related characteristic-value and vector problems, including

1. $\det(A - \lambda I) = 0$, where A is a real symmetric matrix;
2. $\det(AB - \lambda I) = 0$ or $\det(A - \lambda B) = 0$, where either A or B is positive definite and both are real symmetric matrices;
3. the inverse characteristic-value problem:[5] given a real symmetric matrix A, find a diagonal matrix D^2 such that DAD has prescribed proper values.

---

[4] University Relations Division, ORINS.

[5] A. C. Downing, Jr., and A. S. Householder, *J. Assoc. Computing Mach.* 3, 203 (1956).

The Givens[6] procedure for computing roots and vectors was applied. Primary limitations of matrix order are probably time and generated error; mechanically, matrices up to order 400 can be handled automatically.

MATRIX INVERSION AND LINEAR EQUATION SOLVER

Origin: A. S. Householder

Participating Member of Panel: F. J. Witt

Background and Status. – This program uses the method outlined by Householder.[7] When completed, it will find the inverse of a matrix of maximum order 128 and will solve any set of such systems of equations. Error estimates will also be computed. At present, the inverse of a matrix of order 19 may be computed in the internal memory. Additional coding is being done, and the effect of round-off error is being investigated. This program uses the (8,32) floating-point system.

SORTING ROUTINES

Participating Member of Panel: J. H. Vander Sluis

Background and Status. – The method for obtaining a monotonically ordered sequence described by Goldstine and von Neumann[8] has been programed in several subroutines for the Oracle. In these routines, complexes[9] of information are ordered by sorting on the principal[9] number in the complex. A routine using the same method, but for which the complex contains only the principal number, was previously written by R. R. Bate and R. R. Coveyou.

Twelve subroutines have been written which will sort according to fixed-point numbers, (8,32) floating-point numbers, and (40,40) floating-point numbers, or according to their respective magnitudes. Six of these routines order sequences which can be fitted into the fast memory (internal sorting), and six order sequences which are of such length as to require magnetic-tape storage (external sorting). The external routines use the corresponding internal routines. If the programmer has occasion to use more than one of these subroutines, it is possible to convert a basic sorting routine to any of the other forms by inserting a few corrections.

The following definitions are necessary in order to describe the routines:

1. S is the spacing or the number of words in the complex; the spacing S is the basic unit and is always moved as a unit;

---

[6] W. Givens, *Numerical Computation of the Characteristic Values of a Real Symmetric Matrix*, ORNL-1574 (March 3, 1954).

[7] A. S. Householder, *On Solving Linear Algebraic Systems*, ORNL-1785 (Sept. 21, 1954).

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

[9] For definitions see ref 8.

2. P is the position of the sort word in the spacing S;
3. N is the total number of words to be sorted;
4. N_b is the number of words to be sorted on each block of magnetic tape, with the possible exception of the final block;
5. N_{bf} is the number of words to be sorted on the final block of magnetic tape;
6. L is the number of words to be sorted in the internal memory and, in the case of the external sort, will be equal to N_b or N_{bf};
7. t is the total number of blocks of magnetic tape required to store the sequence;
8. adding a prime to a symbol L defines the number $L' = 2^n$, such that

$$2^{n-1} < L \leqq 2^n \ .$$

For the internal sort, the standard compiler subroutine entry is used, and two linkage words are required. These linkage words are

$$\sim \mathrm{M}(P) \qquad \sim \mathrm{M}(L) \ ,$$
$$\sim \mathrm{M}(A_i) \qquad \sim \mathrm{M}(S) \ ,$$

where $\mathrm{M}(A_i) = \mathrm{M}_0(A) + (L' - 1)S$, and P, L, S, and L' are defined above. The beginning memory position of the sequence to be sorted is $\mathrm{M}(A_i)$, and $\mathrm{M}_0(A)$ is the beginning memory position of the ordered sequence. The memory space required for this operation is $(L' + L - 1)S$ consecutive words, starting with $\mathrm{M}_0(A)$. The length of the sequence which can be ordered internally is thus determined by the above space requirement. The programmer must reserve as much consecutive space as is necessary for the sorting process.

The external-sort routines use the standard compiler subroutine entry and require two linkage words, which are

$$0\ PP\ SS\ 00\ D_0 D_1 D_2 \ ,$$
$$\sim \mathrm{M}(N_b) \qquad \sim \mathrm{M}(N) \ ,$$

where P, S, N_b, and N are defined above. The hex character D_0 specifies the magnetic-tape drive on which the unordered data are stored and on which the ordered data will be found; D_1 and D_2 signify the drives which will be used as intermediate storage during the sorting operation. The only limitation on the length of the sequence which can be ordered by the external-sort routine is the number of magnetic-tape blocks available on each of the drives. Three blocks of the internal memory are set aside by the external sort as temporary storages for the ordering process. These are the zero hunks F39, F3A, and F3B.

Representative time estimates for the external routines are given in Table 1. Internal sorting of the blocks of data required approximately one-half the time spent in case 2, the remainder of the time being required to mesh the sorted blocks. The time

Table 1. Representative Time Estimates for External Routines

	Case 1	Case 2
N	340	4096
Fixed, sec	42	166
Fixed magnitude, sec	42	147
(8,32), sec	48	277
(8,32) magnitude, sec	46	244
(40,40), sec	47	364
(40,40) magnitude, sec	46	300

estimates given are maximum times, since generated random numbers were used in the tests.

For case 1, the parameters are $S = 7$, $P = 3$, $N_b = 18$, $N = 340$. For case 2, the parameters for all but (40,40) are $S = 1$, $P = 1$, $N_b = 128$, $N = 4096$. For case 2, (40,40), the parameters are $S = 2$, $P = 1$, $N_b = 64$, $N = 4096$.

VECTOR-BLOCK MATRIX CODES

Origin: J. W. Givens

Participating Member of Panel: M. R. Arnette

References: Mathematics Panel Semiannual Progress Reports, ORNL-2037, -2134

Background and Status. – The complex of codes which will obtain the solution of several systems of linear equations, as well as calculate the inverse of a square matrix, has been revised for the change to 128 words per block on magnetic tape. These codes have now been checked on the Oracle. The input and output codes are being rewritten to take advantage of the new input-output equipment, and the whole system is being unified. The input is to be such that a key word, supplied by the user, will indicate whether the elements are to be found as packed floating-point decimal numbers, or as binary numbers that are in fixed point, (8,32) floating point, or (28,12) floating point. The output code will combine all the exponents and will either leave each element with one exponent or leave all the elements of each column with a single exponent. The details of the output are not yet complete.

ADDITIONS TO THE SUBROUTINE LIBRARY

Participating Member of Panel: B. J. Osborne

Reference: "Sorting Routines," this report

Status. – Several new routines have been received for the subroutine library but have not yet been placed on the compiler. These consist of a curve-plotter code edit, curve-plotter post-mortem, fixed-point arc-tangent, code expansion and contraction, two codes

for finding references to given addresses, and a group of sorting routines. The (8,32) floating-point square root has been rewritten and is available on the compiler. The first two routines are located at the console.

A complete writeup for the compiler has been prepared. It describes in greater detail some of the features that proved most difficult to understand in the last writeup. It also includes some of the features of the new compiler, which is now ready for general use.

The new compiler is a revision of the present one. New features are as follows:

1. *Intersegment storage.* – This feature provides storage for a program that is too large to fit in the memory at one time. Each memory load will be called a "segment" and will be compiled around the storage space used collectively by the segments.

2. *Ability to place the main hunk anywhere.* – This means that a programmer can compile a program to fit in with other programs by specifying the beginning memory address.

3. *A record of where each hunk is called for.* – Previously, if the compiler stopped because it could not find a hunk that was referred to, the programmer had no record of where this reference occurred. Now the compiler will keep a list of each hunk referred to, together with the hunk containing that reference.

4. *Placement of the F00 hunk.* – The F00 hunk will now be placed immediately following the constants rather than at the end of the memory. This means that the whole memory does not have to be punched in order to get a complete copy of the program if the program is less than a memory load.

5. *Read-in blocks.* – The compiler has a number of zero read-in blocks for programmers to use that will leave vacant in the memory the number of cells called for by the program. These are now placed at the end of the memory so that they do not have to be punched out in obtaining a paper-tape copy of the compiled code.

The service routines at the Oracle console have been tested and completely written up so that the operator can now help the programmer with any of these. They include the three curve-plotter codes, tape-to-tape conversions, 10.11*a* and 10.11*b* inputs, post-mortem, tape-to-tape comparison, correction routine, and tape reproducer.

STATISTICS

DISEASE-INCIDENCE ESTIMATION IN POPULATIONS SUBJECT TO MULTIPLE CAUSES OF DEATH

Origin: A. C. Upton, Biology Division

Participating Member of Panel: A. W. Kimball

Reference: Mathematics Panel Semiannual Progress Report, ORNL-2134

Background. – This problem and one solution to it were described in the reference, p 13–15. The solution required the assumption of specific probability distributions for times of death in the population. In addition, the estimation procedures involved lengthy computations requiring several hours to complete. As a result of a search for a simpler solution, the method about to be described was found and has proved more satisfactory. It is essentially distribution free, and the computation procedures are elementary.

Status. – A random sample of N animals from an infinite population of animals is observed until all animals in the sample have died. At the time of death, the cause of death and age of each animal are recorded. The observation period is divided into n intervals, not necessarily of equal length, and the number of animals dying from each cause of death is calculated for each time interval.

Let

$$a_{ij} = \text{number of animals that died of cause } i \text{ during the } j\text{th interval} \quad (i = 1, \ldots, m;\ j = 1, \ldots, n) ,$$

$$A_i = \sum_{j=1}^{n} a_{ij} = \text{total number of animals that died of cause } i ,$$

$$r_j = \sum_{i=1}^{m} a_{ij} = \text{total number of deaths in the } j\text{th interval} .$$

Then

$$\sum_{i=1}^{m} A_i = \sum_{j=1}^{n} r_j = N .$$

The conditional probability, p_{ij}, is defined to be *the probability of death in the jth interval from the ith cause, given survival through the first (j – 1) intervals*. Then

$$s_j = \sum_{i=1}^{m} p_{ij} = \text{probability of death in the } j\text{th interval, given survival through the first } (j - 1) \text{ intervals} ,$$

and it will be convenient to define $s_0 = 0$. The nth interval can be taken as unbounded on the right, whereupon $s_n = 1$. From the definition of the p_{ij},

$$R_j = (1 - s_0)(1 - s_1) \ldots (1 - s_{j-1})$$

$$= \text{probability of survival through the first } (j - 1) \text{ intervals} .$$

Thus $p_{ij}R_j$ is *the unconditional probability of death in the* j*th interval from the* i*th cause*. It may be verified that

$$\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} R_j = 1 .$$

For the sample of N animals, the likelihood function is

$$\frac{N!}{\prod_{i=1}^{m} \prod_{j=1}^{n} (a_{ij}!)} \prod_{i=1}^{m} \prod_{j=1}^{n} \left(p_{ij} R_j \right)^{a_{ij}} ,$$

and its logarithm may be written

$$L = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \log p_{ij} + \sum_{j=1}^{n} S_j \log (1 - s_{j-1}) ,$$

where

$S_j = r_j + r_{j+1} + \ldots + r_n$ = number of animals surviving at the beginning of the jth interval.

The maximum-likelihood estimates of the p_{ij} are

$$\hat{p}_{ij} = \frac{a_{ij}}{S_j} ,$$

and the estimate of the disease incidences when all causes of death are operating is

$$\hat{I}_i = \sum_{j=1}^{n} \hat{p}_{ij} \hat{R}_j = \frac{A_i}{N} .$$

A more difficult problem arises when it is desired to estimate the disease incidences in populations in which one or more causes of death have been eliminated. If the first cause of death is eliminated, the new conditional probabilities are

$$p'_{ij} = \frac{p_{ij}}{1 - p_{1j}} \qquad (i = 2, \ldots, m;\ j = 1, \ldots, n) ,$$

and the true incidence of the ith disease is now

$$I'_i = \sum_{j=1}^{n} p'_{ij} R'_j \qquad (i = 2, \ldots, m) ,$$

where

$$R'_j = (1 - q_0)(1 - q_1) \ldots (1 - q_{j-1}) ,$$

$$q_j = p'_{2j} + \ldots + p'_{mj} ,$$

$$q_0 = 0 .$$

Then

$$\hat{I}_i' = \sum_{j=1}^{n} \hat{p}_{ij}' \hat{R}_j'$$

is the maximum-likelihood estimate of I_i'. In reduced form suitable for calculation,

$$\hat{I}_i' = \frac{1}{N} \sum_{j=1}^{n} \frac{a_{ij}}{(1 - \hat{p}_{11})(1 - \hat{p}_{12}) \ldots (1 - \hat{p}_{1j})} .$$

An approximate variance for $\hat{I}_i'$ is given by

$$V(\hat{I}_i') = \frac{1}{N^3} \sum_{j=1}^{n} \frac{a_{ij}}{\left[(1 - \hat{p}_{11})(1 - \hat{p}_{12}) \ldots (1 - \hat{p}_{1j})\right]^2} .$$

The sampling properties of the $\hat{I}_i'$ are being investigated by means of random-sampling experiments. This study should provide some evidence concerning the effects of sample size, choice of time intervals, and other variables on the bias of the estimates and on the validity of the approximate variance formula. The method described in this report is the subject of a more extensive paper which will be submitted for publication in the open literature.

INVERSE OF A CERTAIN PATTERNED MATRIX

Origin: A. W. Kimball

Participating Member of Panel: M. A. Kastenbaum

Background and Status. – The matrix of asymptotic variances and covariances of the estimates of the parameters of a multinomial distribution may be found by inverting the following symmetric positive definite matrix:

$$\left(-\mathcal{E} \frac{\partial^2 L}{\partial p_i \partial p_{i'}}\right) , \tag{1}$$

where $\mathcal{E}$ is the operator of mathematical expectation, L is the logarithm of the multinomial density function, and the p_i are parameters of the distribution, for i and $i' = 1, 2, \ldots, k$.

In most instances the elements of matrix 1 follow a very clear pattern, and may therefore be inverted without too much difficulty by applying the rather elegant techniques discussed by Roy and Sarhan.[1]

[1] S. N. Roy and A. E. Sarhan, *Biometrika* 43, Parts 1 and 2, p 227 (1956).

Recently, in a problem involving the estimation of the parameters of a distribution very similar to the multinomial distribution, it became necessary to evaluate the asymptotic variance-covariance matrix as well. In this instance the inversion of the following $k \times k$ matrix, derived from matrix 1, presented some difficulty:

$$(2)\quad B(k \times k) = \begin{pmatrix} \frac{1}{p_1} + \frac{(1-c)}{(1-t)} & \frac{(1-c)}{(1-t)} & \frac{(1-c)}{(1-t)} & \cdots & \frac{(1-c)}{(1-t)} \\ \frac{(1-c)}{(1-t)} & \frac{1}{p_2} + \frac{(1-c)}{(1-t)} & \frac{(1-c)}{(1-t)} & \cdots & \frac{(1-c)}{(1-t)} \\ \frac{(1-c)}{(1-t)} & \frac{(1-c)}{(1-t)} & \frac{1}{p_3} + \frac{(1-c)}{(1-t)} & \cdots & \frac{(1-c)}{(1-t)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{(1-c)}{(1-t)} & \frac{(1-c)}{(1-t)} & \frac{(1-c)}{(1-t)} & \cdots & \frac{1}{p_k} + \frac{(1-c)}{(1-t)} \end{pmatrix},$$

where

$$t = \sum_{i=1}^{k} p_i \ .$$

The problem was to find B^{-1}, the inverse of B.

Roy and Sarhan[1] prove the following theorem: If

$$(3)\qquad A = D_{1/p_i} - \frac{\lambda}{1 + \lambda \sum_{i=1}^{k} q_i \frac{r_i}{p_i}} \overrightarrow{(q/p)}\, \overrightarrow{(r/p)}^T \ ,$$

then

$$(4)\qquad A^{-1} = D_{p_i} + \lambda \vec{(q)}\, \vec{(r)}^T \ ,$$

where

A is a $k \times k$ matrix, and A^{-1} its inverse,

D_{1/p_i} is a $k \times k$ diagonal matrix with elements $(1/p_1, 1/p_2, \ldots, 1/p_k)$,

D_{p_i} is a $k \times k$ diagonal matrix with elements $(p_1, p_2, \ldots, p_k)$,

$\overrightarrow{(q/p)}$ is the $k \times 1$ vector $\begin{pmatrix} q_1/p_1 \\ q_2/p_2 \\ \cdot \\ \cdot \\ \cdot \\ q_k/p_k \end{pmatrix}$,

$\overrightarrow{(r/p)}^T$ is the $1 \times k$ vector $(r_1/p_1, r_2/p_2, \ldots, r_k/p_k)$,

$(\vec{q})$ is the $k \times 1$ vector $\begin{pmatrix} q_1 \\ q_2 \\ \cdot \\ \cdot \\ \cdot \\ q_k \end{pmatrix}$,

$(\vec{r})^T$ is the $1 \times k$ vector $(r_1, r_2, \ldots, r_k)$,

λ and $\sum_{i=1}^{k} q_i \dfrac{r_i}{p_i}$ are scalars.

If Eq. 2 is now rewritten as

$$B = D_{1/p_i} + \frac{1 - c}{1 - t} J , \tag{5}$$

where J is a $k \times k$ matrix, all of whose elements are unity, and if, in Eq. 3,

$$\lambda = - \frac{(1 - c)}{(1 - ct)} \text{ and } (\vec{q}) = (\vec{r}) = (\vec{p}) ,$$

then Eq. 3 becomes

$$A = D_{1/p_i} + \frac{1 - c}{1 - t} J = B . \tag{6}$$

It therefore follows, from Eq. 4, that

$$B^{-1} = D_{p_i} - \frac{(1 - c)}{(1 - ct)} (\vec{p})(\vec{p})^T . \tag{7}$$

To show that $BB^{-1} = I$, the identity matrix, write

$$BB^{-1} = D_{1/p_i} D_{p_i} + \frac{(1 - c)}{(1 - t)} J D_{p_i} - \frac{(1 - c)}{(1 - ct)} D_{1/p_i} (\vec{p})(\vec{p})^T - \tag{8}$$

$$- \frac{(1 - c)(1 - c)}{(1 - t)(1 - ct)} J (\vec{p})(\vec{p})^T .$$

Note that

$$D_{1/p_i} D_{p_i} = I(k \times k)$$

and that

$$(9) \quad JD_{p_i} = D_{1/p_i} (\vec{p})(\vec{p})^T = \frac{1}{t} J(\vec{p})(\vec{p})^T$$

$$= \begin{pmatrix} p_1 & p_2 & p_3 & \cdots & p_k \\ p_1 & p_2 & p_3 & \cdots & p_k \\ p_1 & p_2 & p_3 & \cdots & p_k \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p_1 & p_2 & p_3 & \cdots & p_k \end{pmatrix} = M(k \times k) .$$

Thus

$$(10) \qquad BB^{-1} = I + \left[\frac{(1-c)}{(1-t)} - \frac{(1-c)}{(1-ct)} - \frac{t(1-c)^2}{(1-t)(1-ct)} \right] M = I ,$$

since the coefficient of M is a scalar which is identically equal to zero. Equation 7 may be rewritten as

$$(11) \quad B^{-1} = \begin{pmatrix} p_1 - p_1^2 \frac{(1-c)}{(1-ct)} & -p_1 p_2 \frac{(1-c)}{(1-ct)} & \cdots & -p_1 p_k \frac{(1-c)}{(1-ct)} \\ -p_1 p_2 \frac{(1-c)}{(1-ct)} & p_2 - p_2^2 \frac{(1-c)}{(1-ct)} & \cdots & -p_2 p_k \frac{(1-c)}{(1-ct)} \\ \cdot & \cdot & \cdot & \cdot \\ -p_1 p_k \frac{(1-c)}{(1-ct)} & -p_2 p_k \frac{(1-c)}{(1-ct)} & \cdots & p_k - p_k^2 \frac{(1-c)}{(1-ct)} \end{pmatrix} .$$

A PROBLEM IN ESTIMATION BY THE METHOD OF MAXIMUM LIKELIHOOD. I

Origin: E. Novitski, Biology Division

Participating Members of Panel: M. A. Kastenbaum, A. W. Kimball

Background and Status. – Male *Drosophila melanogaster* carrying a translocation between the X chromosome and the minute fourth chromosome, as well as a Y chromosome and a normal fourth chromosome, are capable of producing four distinct types of gametes which shall be identified as AB, A′B′, AB′, A′B. When such males are mated to attached-X females with a Y chromosome, AB, A′B′, and AB′ are recoverable, and

A′B is lethal. If the Y chromosome of the female is replaced by the proximal segment of the translocation, the recoverable gamete types are A′B′, A′B, and AB′, with AB semilethal or lethal. The problem is to estimate the relative frequencies of the four gamete types.

If the true probabilities of observing the four types of gametes are p_1 for AB, p_2 for A′B′, p_3 for A′B, and p_4 for AB′, where

$$p_1 + p_2 + p_3 + p_4 = 1 , \tag{1}$$

then, in experiments of the first type, where A′B cannot be observed, the probabilities associated with the three recoverable types of gametes may be specified as

$$\begin{cases} \theta_1 = \dfrac{p_1}{1 - p_3} , & \text{for AB} , \\[2ex] \theta_2 = \dfrac{p_2}{1 - p_3} , & \text{for A′B′} , \\[2ex] \theta_4 = \dfrac{p_4}{1 - p_3} , & \text{for AB′} . \end{cases} \tag{2}$$

Similarly, in experiments of the second type, the probabilities may be specified as

$$\begin{cases} \phi_2 = \dfrac{p_2}{1 - p_1} , & \text{for A′B′} , \\[2ex] \phi_3 = \dfrac{p_3}{1 - p_1} , & \text{for A′B} , \\[2ex] \phi_4 = \dfrac{p_4}{1 - p_1} , & \text{for AB′} . \end{cases} \tag{3}$$

Then, in the first experiment, the probability of observing a_1 gametes of type AB, a_2 of type A′B′, and a_4 of type AB′, out of a sample of size N_1, is given by

$$P_1 = \frac{N_1!}{a_1! a_2! a_4!} \theta_1^{a_1} \theta_2^{a_2} \theta_4^{a_4} , \tag{4}$$

where

$a_1 + a_2 + a_4 = N_1$, and $\theta_1 + \theta_2 + \theta_4 = 1$.

Similarly, the probability of observing b_2 gametes of type A′B′, b_3 of type A′B, and b_4 of type AB′, out of a sample of size N_2 from the second type of experiment, is

$$P_2 = \frac{N_2!}{b_2! b_3! b_4!} \phi_2^{b_2} \phi_3^{b_3} \phi_4^{b_4} , \tag{5}$$

where

$$b_2 + b_3 + b_4 = N_2, \text{ and } \phi_2 + \phi_3 + \phi_4 = 1.$$

The probability of observing these results simultaneously in two independent experiments is given by the product of Eqs. 4 and 5 as follows:

$$\omega = P_1 P_2 = \frac{N_1! N_2!}{a_1! a_2! a_4! b_2! b_3! b_4!} \theta_1^{a_1} \theta_2^{a_2} \theta_4^{a_4} \phi_2^{b_2} \phi_3^{b_3} \phi_4^{b_4} . \tag{6}$$

Equation 6, taken together with the relationships defined in Eqs. 1, 2, and 3, becomes

$$\omega = \frac{N_1! N_2!}{a_1! a_2! a_4! b_2! b_3! b_4!} p_1^{a_1} p_2^{(a_2+b_2)} \times \tag{7}$$

$$\times\, p_3^{b_3} (1 - p_1 - p_2 - p_3)^{(a_4+b_4)} (1 - p_3)^{-N_1} (1 - p_1)^{-N_2} .$$

Taking the logarithm of Eq. 7 gives

$$L = \log \omega = \text{const} + a_1 \log p_1 + (a_2 + b_2) \log p_2 + b_3 \log p_3 + \tag{8}$$

$$+ (a_4 + b_4) \log (1 - p_1 - p_2 - p_3) - N_1 \log (1 - p_3) - N_2 \log (1 - p_1) .$$

To estimate the parameters of the distribution, Eq. 8 is maximized as follows:

$$\frac{\partial L}{\partial p_1} = \frac{a_1}{p_1} + \frac{N_2}{(1 - p_1)} - \frac{(a_4 + b_4)}{(1 - p_1 - p_2 - p_3)} = 0 , \tag{9}$$

$$\frac{\partial L}{\partial p_2} = \frac{a_2 + b_2}{p_2} - \frac{(a_4 + b_4)}{(1 - p_1 - p_2 - p_3)} = 0 , \tag{10}$$

$$\frac{\partial L}{\partial p_3} = \frac{b_3}{p_3} + \frac{N_1}{(1 - p_3)} - \frac{(a_4 + b_4)}{(1 - p_1 - p_2 - p_3)} = 0 . \tag{11}$$

The solution of these three simultaneous equations will yield maximum-likelihood estimates of p_1, p_2, and p_3, and, ultimately, p_4.

From Eq. 10 it can be shown that

$$p_2 = \frac{(a_2 + b_2)(1 - p_1 - p_3)}{(a_2 + b_2 + a_4 + b_4)} ; \tag{12}$$

hence, upon substitution, Eqs. 9 and 11 become, respectively,

$$\frac{a_1}{p_1} + \frac{N_2}{(1 - p_1)} - \frac{(a_2 + b_2 + a_4 + b_4)}{(1 - p_1 - p_3)} = 0 \tag{13}$$

and

$$\frac{b_3}{p_3} + \frac{N_1}{(1 - p_3)} - \frac{(a_2 + b_2 + a_4 + b_4)}{(1 - p_1 - p_3)} = 0 . \tag{14}$$

The only admissible solution of Eqs. 13 and 14 is

$$\hat{p}_1 = \frac{a_1(b_2 + b_4)}{b_3(a_2 + a_4) + N_1(b_2 + b_4)} . \tag{15}$$

and

$$\hat{p}_3 = \frac{b_3(a_2 + a_4)}{b_3(a_2 + a_4) + N_1(b_2 + b_4)} . \tag{16}$$

Substituting these values in Eq. 12 gives

$$\hat{p}_2 = \frac{(a_2 + b_2)}{(a_2 + b_2 + a_4 + b_4)} \left[\frac{(N_1 - a_1)(b_2 + b_4)}{b_3(a_2 + a_4) + N_1(b_2 + b_4)} \right] . \tag{17}$$

Finally,

$$\hat{p}_4 = 1 - \hat{p}_1 - \hat{p}_2 - \hat{p}_3 . \tag{18}$$

The variances and covariances of these estimates have not been evaluated as yet.

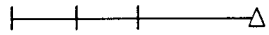
A PROBLEM IN ESTIMATION BY THE METHOD OF MAXIMUM LIKELIHOOD. II

Origin: A. C. Fabergé, Biology Division

Participating Member of Panel: M. A. Kastenbaum

Background and Status. – Let a maize endosperm chromosome arm carrying two dominant markers A and B be represented as

E A B SFA



where E stands for the natural stable end of the chromosome arm and SFA for the spindle-fiber attachment. For such an arm, five classes of offspring, effected by irradiation, are recognizable:

1. Cy(EA) , 2. Cy(AB) , 3. L(A) , 4. L(B) , 5. L(AB) ,

where Cy represents a breakage-fusion-bridge cycle, and L is the loss of a marker.

If the probability that an offspring falls into class i (for $i = 1, 2, 3, 4, 5$) is given by p_i, then the probability that, in a sample of size N, n_i offspring are of class i, for all i, is given by the multinomial density function

$$\phi = \frac{N!}{\prod_{i=1}^{5} n_i!} \prod_{i=1}^{5} p_i^{n_i} , \tag{1}$$

where

$$\sum_{i=1}^{5} p_i = 1, \text{ and } \sum_{i=1}^{5} n_i = N.$$

The problem is to find estimates of the parameters p_i, for a given set of n_i and N, subject to the additional constraint that

$$p_5 = p_1 p_4 \left(\frac{1}{p_2} + \frac{1}{p_3}\right) . \tag{2}$$

Using the method of maximum likelihood, Eq. 1 may be maximized, subject to the constraint imposed in Eq. 2 and subject to the further constraint that

$$\sum_{i=1}^{5} p_i = 1 .$$

This procedure leads to the following four simultaneous equations in four of the unknowns:

$$(n_1 + n_5) - Np_1 \left[1 + p_4 \left(\frac{1}{p_2} + \frac{1}{p_3}\right)\right] = 0 , \tag{3}$$

$$(n_4 + n_5) - Np_4 \left[1 + p_1 \left(\frac{1}{p_2} + \frac{1}{p_3}\right)\right] = 0 , \tag{4}$$

$$n_5 - \left(\frac{1}{p_2} + \frac{1}{p_3}\right)[(n_2 - Np_2)p_2 + Np_1 p_4] = 0 , \tag{5}$$

$$n_5 - \left(\frac{1}{p_2} + \frac{1}{p_3}\right)[(n_3 - Np_3)p_3 + Np_1 p_4] = 0 . \tag{6}$$

For a given set of n_i and N, the maximum-likelihood estimates, $\hat{p}_i$ (i = 1, 2, 3, 4), may be evaluated by an iterative solution of these equations, where the "best first guess" of p_2 is n_2/N, of p_3 is n_3/N, and of p_4 is n_4/N.

A NONLINEAR DOSE-EFFECT FITTING PROBLEM

Origin: L. M. Sandler, Biology Division

Participating Members of Panel: A. W. Kimball, E. Leach

Background and Status. – The equation

$$y = \frac{1}{e^m - 1}(e^{mbx} - 1)$$

was proposed as a theoretical relationship between the proportion of recessive lethals (y) and x-ray dose (x) in populations of *Drosophila melanogaster*. A least-squares fit to experimental data was desired.

To obtain initial estimates of the parameters m and b, the quadratic approximation

$$y = \frac{1}{e^m - 1}\left[mbx + \frac{(mbx)^2}{2}\right] = q_1 x + q_2 x^2$$

was used. In this expression

$$q_1 = \frac{mb}{e^m - 1}, \qquad q_2 = \frac{m^2 b^2}{2(e^m - 1)},$$

and, after least-squares estimates of q_1 and q_2 were obtained by conventional methods, initial estimates of m and b were given by

$$m_0 = \log_e \frac{q_1^2 + 2q_2}{q_1^2}, \qquad b_0 = \frac{q_1(e^{m_0} - 1)}{m_0}.$$

The original equation can be written in the form

$$u = bx,$$

where

$$u = \frac{1}{m}\log_e [y(e^m - 1) + 1].$$

Final estimates of the parameters were obtained by minimizing

$$V = \Sigma(u - bx)^2,$$

the minimization with respect to m being done graphically in the neighborhood of m_0; that is, trial values of m near m_0 were chosen, and V, minimized with respect to b, was plotted against the trial values of m. Five points were found to be sufficient to determine $\hat{b}$ and $\hat{m}$ to the accuracy required.

ESTIMATION OF THE DIAMETER OF THE *HABROBRACON* NUCLEUS

Origin: R. C. von Borstel, Biology Division

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

Background and Status. – When *Habrobracon* eggs are exposed to alpha particles directed at the side on which the nucleus is located, the survival curve takes the familiar exponential form,

$$p = e^{-bx},$$

where p is the surviving fraction and x is the length of exposure in seconds. When the mean number of hits per nucleus is 1, the expected frequency of the class receiving no hits is $1/e$. Thus the dose D corresponding to a fraction $1/e$ surviving is the dose which yields a mean of one hit per nucleus; that is,

$$\log_e \frac{1}{e} = -bD,$$

or

$$\frac{1}{D} = b \ .$$

From a least-squares fit of the survival data, the estimate obtained was

$$\frac{1}{D} = 15.652 \times 10^{-3} \text{ alpha particles/sec} \ ,$$

with a variance of 2.2247×10^{-6}.

In an independent experiment using single-crystal zinc sulfide scintillation targets of known dimensions with a polonium-silver radiation source at the same distance, it was determined that the number of alpha particles hitting a target of 1 μ^2 in area was 3.3229×10^{-3}/sec, with a variance of 0.007367×10^{-6}. Thus an estimate of the area of the egg nucleus, seen as a target by the source, is given by

$$A = \frac{15.652 \times 10^{-3}}{3.3229 \times 10^{-3}} \mu^2 = 4.7103 \mu^2 \ ,$$

and this yields a diameter $d = 2.45\ \mu$.

The variance of A, computed from the usual formula for the variance of a ratio, is 0.21628. Since

$$d = \frac{2}{\sqrt{\pi}} \sqrt{A} \ ,$$

the variance of d is approximately $V(A)/\pi A$. This yields a standard error for d of $0.12\ \mu$.

PROBLEMS

APPLICATION OF STANDARD CODES

Participating Member of Panel: B. J. Osborne

Origin: S. I. Cohen, Reactor Experimental Engineering Division

Problem. – A least-squares fit of several sets of experimental data to a polynomial was obtained by using the linear-regression code.

Origin: T. Gens, Chemistry Division

Problem. – A code for the calculation of energies of activation was written, and several cases have been run. It is expected that a large number will be run in the near future.

Origin: R. M. Healy, Chemistry Division

Problem. – A least-squares fit was obtained for

$$\ln P = \frac{-a}{t} + b \ln t + cx ,$$

given P, t, and x.

Origin: H. W. Morgan, Stable Isotopes Division

Problem. – A least-squares fit of experimental data to the polynomial

$$y = ax + bx^2$$

was obtained. Several other sets of data were fitted to polynomials of order 4 and 5. In addition, eight systems of simultaneous linear equations of order 15 were solved.

Origin: M. H. Lietzke, Chemistry Division

Problem. – Several sets of experimental data were fitted, in the sense of least-square residuals, to polynomials of the form

$$s = s_0 + s_1 m + s_2 m^2 + s_3 m^3 + s_4 m^4 .$$

Origin: H. L. Yakel, Jr., Health Physics Division

Problem. – Least-squares solutions to a polynomial of order 3 were obtained for seven sets of data.

Origin: G. J. Atta, Mathematics Panel

Problem. – The inverses of several matrices of orders up to 10 were obtained.

Origin: L. Holway, Raytheon Manufacturing Company

Problem. – Several sets of parameters were used in the two-dimensional reactor code.[1]

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

Problem. – The code for the calculation of stresses in piping systems has been used regularly during this period.

SHORT-LIVED FISSION-PRODUCT GAMMA RAYS

Origin: W. Zobel, Applied Nuclear Physics Division

Participating Member of Panel: F. J. Witt

Background. – From a fissioning nucleus there are emitted particles and electromagnetic radiation. The gammas emitted immediately upon fission are the prompt fission gamma rays. The fissioning nucleus usually breaks up into two parts which are called the fission products. These parts, in turn, emit particles and gamma rays. The gamma rays thus emitted are the fission-product gamma rays.

---

[1] C. L. Bradshaw, *A Three-Region, Two-Group, Two-Dimensional Reactor Code for the Oracle*, ORNL-2047 (March 15, 1956).

Part I. Measurement of Energy Spectra at Discrete Times After Fission. – This part of the experiment was designed to measure the gross fission-product gamma-ray spectrum as a function of energy and time after fission (for short times after fission). Data were taken from both pair and Compton spectrometers.

For the Compton spectrometer, let N be defined by

$$N = \frac{\frac{N_p}{\Delta P} \cdot C(N) \cdot \frac{\Delta E_p}{\Delta E_\gamma}}{S \cdot \frac{\Delta E_p}{\Delta P} \cdot t \cdot (\epsilon\omega)} .$$

Here $N_p/\Delta P$ is an experimental number for the Compton spectrometer determined by the voltage channel; $C(N)$ is the count loss correction; $\Delta E_p/\Delta E_\gamma$ is a conversion factor from pulse-height channel width to energy channel width and is given by

$$\frac{\Delta E_p}{\Delta E_\gamma} = \frac{(1 + \sigma E_\gamma)^2 - 1}{(1 + \sigma E_\gamma)^2} ,$$

where E_γ is the energy of the incident gamma ray and equals

$$P \cdot \frac{\Delta E}{\Delta P} \cdot \frac{E_\gamma}{E_p} ,$$

where P is the channel number and E_γ/E_p is the ratio of gamma energy to corresponding pulse height; $\Delta E_p/\Delta P$ is a pulse-height-to-energy conversion factor. The symbol σ is a constant of equivalence, S is the source term, $(\epsilon\omega)$ is the efficiency of the spectrometer, and t is the counting time.

For the pair spectrometer,

$$N = \frac{\frac{N_p}{\Delta P} \cdot C(N)}{S \cdot \left(\frac{\Delta E_p}{\Delta P}\right) \cdot t \cdot (\epsilon\omega)} ,$$

where

$$E_\gamma = P \cdot \frac{\Delta E}{\Delta P} + 1.022.$$

The values E_γ, N, and the standard error of N were calculated for available data from 60-channel analyzers, together with the positive area under the curve (E_γ vs N) for each set of data.

This code was written with sufficient generality so as to be usable in other experiments.

Part II. Measurement of Time Spectra of Large Energy Groups. – In the second part of the experiment, the detailed time dependence of relatively large photon groups was studied. In this case, N is defined as follows:

$$N = \frac{\frac{N_p}{\Delta P}\, C(N)}{\tau \cdot S \cdot \Delta E \cdot (\epsilon\omega)} ,$$

where $N_p/\Delta P$, $C(N)$, S, and $(\epsilon\omega)$ are defined as in Part I, ΔE is the energy interval within which gamma rays are counted, τ is the average counting time per channel and is equal to

$$\frac{2}{15}\left(\frac{R_1}{m_1} + \frac{R_2}{m_2}\right) ,$$

where m_1 is the number of sweeps in the time calculation before the experiment starts, m_2 is the number of sweeps at the end, R_1 is the number of calculated counts in each channel before the experiment starts, and R_2 is the number of counts at the end. If

$$a = \frac{2}{15}\frac{R_1}{m_1} , \qquad \bar{a} = \frac{2}{15}\frac{R_2}{m_2} ,$$

and δ_1, $\Delta\delta_1$, δ_2, $\Delta\delta_2$ are experimental constants, then the following equalities hold:

$$\begin{aligned}
b_1 &= a_1 + \delta_1 , & \bar{b}_1 &= \bar{a}_1 + \delta_2 ,\\
b_2 &= b_1 + a_1 + a_2 , & \bar{b}_2 &= \bar{b}_1 + \bar{a}_1 + \bar{a}_2 ,\\
&\cdots , & &\cdots ,\\
b_{20} &= b_{19} + a_{19} + a_{20} , & \bar{b}_{20} &= \bar{b}_{19} + \bar{a}_{19} + \bar{a}_{20} ,\\
b_{21} &= b_{20} + \Delta\delta_1 + a_{20} + a_{21} , & \bar{b}_{21} &= \bar{b}_{20} + \bar{a}_{20} + \Delta\delta_2 + \bar{a}_{21} ,\\
&\cdots , & &\cdots ,\\
b_{40} &= b_{39} + a_{39} + a_{40} , & \bar{b}_{40} &= \bar{b}_{39} + \bar{a}_{39} + \bar{a}_{40} .
\end{aligned}$$

Since

$$t' = \frac{b_i + \bar{b}_i}{2} + d ,$$

where d is a correction constant, the count rate is found by interpolating with t' from experimental data, and the count loss correction $C(N)$ is found by interpolation with this result and experimental data. The code computes N and the standard error in N.

ECHELLOGRAM DATA REDUCTION

Origin: K. L. Vander Sluis, Stable Isotopes Division

Participating Member of Panel: J. H. Vander Sluis

References: Mathematics Panel Semiannual Progress Reports, ORNL-2037 and -2134, and "Sorting Routines," this report

Status. – The key to this problem is an Oracle subroutine which can monotonically order complexes according to any number in the complex, both for fixed-point numbers and for (8,32) floating-point numbers. It has been necessary to rewrite some of the previous parts for use in the more general reduction of echellogram data. The sorting routine has been completed and is described in the section mentioned in the reference. Work is progressing on the remainder of the problem.

DETERMINATION OF THE BREIT-WIGNER CONSTANTS FOR XENON

Origin: E. C. Smith, S. Bernstein

Participating Members of Panel: S. E. Atta, N. M. Dismuke

Reference: Mathematics Panel Semiannual Progress Report, ORNL-2134

Status. – Calculation of the total and capture g factors for Xe^{135}, where

$$g(kT) = \frac{2}{\sqrt{\pi}} \sqrt{\frac{kT}{0.0253}} \frac{\overline{\sigma}(kT)}{\sigma(293°K, 0.0253 \text{ ev})},$$

has been completed for T ranging from 100 to 1500°K. Computations with and without Doppler corrections were made.

CHARACTERISTICS OF URANYL SULFATE SOLUTIONS

Origin: G. M. Hebert and D. W. Sherwood, Chemistry Division

Participating Members of Panel: C. D. Griffies, C. L. Gerberich

Background and Status. – Solutions of various volumes and concentrations of uranyl sulfate were placed in fused-quartz tubes and heated. Measurements were taken at various temperatures of the total height of the solution and of the heights of the light and heavy phases. Other data available were the total volume of the solution at room temperature, the density of the solution, and the volumes of the cone-shaped ends of the tube. Calculations were made on the Oracle of (1) the total volume of the tube, (2) the volume of each phase of the solution, and (3) the percentage volume of each phase of the solution in relation to the total volume of the tube.

This problem is now in a production status. A curve-plotter output code will be added.

ORACLE OPERATIONS, SERVICE ROUTINES, AND EDUCATION

OPERATION OF THE ORACLE

From July 1956 through February 1957, a total of 4246 hr of computing time was used by programmers, of which 28% was for "debugging" and 72% was for production.

The Mathematics Panel used 12% of the total computing time for mathematical research and the preparation of subroutines, and 13% was used for solving problems originating outside the Panel. The remaining 75% was used for computations programed by persons who were not members of the Mathematics Panel.

New block counters for the magnetic-tape auxiliary memory have been installed, which record the number of the block under the head in hexadecimal rather than in decimal form. Weighing baskets have also been installed for each drive. Their purpose is to automatically wind or unwind magnetic tape as needed.

New safety features have been added to the magnetic-tape auxiliary memory. Power for all drives is immediately shut off when any tape-drive door is opened, thus eliminating the danger of electrical or mechanical injury to personnel.

CURVE-PLOTTER DUMP

Participating Member of Panel: C. L. Gerberich

Reference: Oracle Memo 13

Background and Status. – It was felt that the curve plotter was not used as much as was expected because of the amount of extra coding needed to program curve-plotter output. Therefore a generalized curve-plotter output was written. This program takes m units of n words each from magnetic tape P and arranges them in 1, 2, 3, or 4 columns on the curve plotter. The code has been written to make it as simple as possible for the user, and the information is supplied to the program in two ways. In order to identify the film, an input tape is used which contains in hexadecimal form the number of characters to be used in the heading. This number must be less than 25. Following this number will be the bi-hex representation of the symbols wanted in the heading, with five bi-hex characters listed per word. At the end of the last word is a double space. By taking Oracle Memo 4 and numbering the list of symbols in hex, starting with 00, the user will have the bi-hex representation used by the program. It is hoped that the user will pick one heading, such as his name, which he will use every time, so that one tape will be sufficient. Any data which will change from time to time are supplied by a second

method. The code stops the machine at two points to receive more information through the Q register. This information includes the number,

m, of units of data,
n, of words per unit,
P, of the magnetic-tape drive containing the information,
c, of columns desired,
M, of the month,
D, of the day,
p, of the starting page.

The information needed at the first stop is

$$Q_1 = P\ xxx\ nnn\ mmm$$

and at the second stop is

$$Q_2 = MM\ DD\ ppp\ xx\ C\ ,$$

where x denotes an irrelevant hexadecimal character. The code then plots the heading supplied by the tape, followed by the date (*MM-DD*-57) and the page number. The page number is advanced in decimal from page to page except for the hundreds digit, which is in hex. Therefore from 999 it advances to A00. The hundreds digit is normally used to identify a particular set of data. Therefore if the user should want to make several runs through this code he could use starting page numbers 000, 100, 200, . . ., A00, . . ., F00.

The program also inserts a block number between the different units. Thus the user can quickly find a particular unit of data. It is assumed that a unit of information must start on a new block on the magnetic tape. Therefore any information on the tape between the end of a unit and the end of the magnetic-tape block is omitted. For a unit of length equal to or less than 128, the unit number is also the block number. The block number is given in hex form.

CURVE-PLOTTER POST-MORTEM

Participating Member of Panel: C. L. Gerberich

Reference: Oracle Memo 12

Background and Status. – The general "debugging" practices still leave a great deal to be desired. It was hoped that a curve-plotter post-mortem might be an aid in improving these methods. The main duty of this program is to copy on the curve plotter three columns of information. Column 1 is the address of the word in the memory. The second column is the word as it was in the memory at the point where the original problem was stopped. The third column is used only when the word in column 2 differs from what it was when the original problem was started. If the two differ, the original word is written

in column 3, if they do not, column 3 is omitted. To use the code the operator copies the original code onto a magnetic tape P_a. Then the deceased code as it remains in the memory is copied onto magnetic tape $P_b (P_a \neq P_b)$. The curve-plotter post-mortem is then applied. The input system for this code works on the same principles as does the curve-plotter dump program. It uses an input tape for the main heading and stops at two points to receive added information through the Q register. This information includes the number,

P_a, of the drive containing the original code,
P_b, of the drive containing the deceased code,
s, indicating the memory size (8 = double, 4 = single),
m, which is the address of the first word to be compared,
n, which is the address of the last word to be compared,
q, which is the address of the first word dumped on tope,
M, of the month,
D, of the day,
p, of the starting page.

Hexadecimal characters not used are indicated by x. The information needed at the first stop is

$$Q_1 = P_a P_b \; s \; x \; mmm \; nnn$$

and at the second stop is

$$Q_2 = MM \; DD \; ppp \; qqq \; .$$

This code is kept in the console with a complete set of operating instructions.

TRACER ROUTINE

Origin: Mathematics Panel

Participating Member of Panel: B. J. Osborne

Background and Status. – A tracer routine is being written for the Oracle which may be used with any program to monitor the execution of any specified order, a list of orders, or all orders. The purpose of the tracer routine is to assist programmers in "debugging" codes.

With the tracer in the memory the following operations take place:

1. A format tape is read in, containing all the necessary information that the tracer will need: the beginning transfer order, whether the code uses single or double memory, the order or orders to be traced, whether the code to be traced is now on paper tape or magnetic tape, how many sections of code there are and the beginning address of each section, the address where the tracer should stop tracing, the type of output wanted, the beginning and end of temporary storage if it is all together, and the number of blocks the

code uses on all four drives. This information will set up the necessary loops in the tracer and will enable it to pedal through the code one step at a time.

2. The code to be traced will then be read in from paper tape or magnetic tape and stored on the tracer tape one block at a time. A list will be made of the addresses contained on all blocks.

3. The tracer will then start pedaling through the code, testing each order to see whether it is to be traced, and if it is, the tracer will record the necessary information, using the type of output specified.

4. The tracer will stop when a stop order is reached, with the choice of starting up again from that spot if the programmer desires. It will stop completely at the address listed on the format tape.

The tracer will work from magnetic tape, which will considerably increase the computing time. This is the price paid for being able to trace orders through a code as long as 2048 words and for being able to trace all orders.

The tracer is being written for the new input and output, and will trace any of the new orders, as well as the ones now being used.

There will be three types of output available: narrow magnetic tape, paper tape, and curve plotter. Paper tape will be practicable only in the case of very few orders being traced. Whichever type is chosen will immediately set up a loop which will block all orders in the code referring to that type of output. This is the only restriction placed on the program.

All drives can be used by the code being traced, since the tracer sets up one drive to be used as all four drives, based upon the information on the format tape. The tracer will keep an inner counter for each drive, as well as an over-all one for the drive being used as all four.

This routine is expected to be completed by the time the Oracle is back in operation following the March shutdown.

ORACLE PROGRAMING CLASS

A preliminary draft of *Notes for Beginning Programmers for the Oracle*, prepared by S. E. Atta, was used in the two-week programing course given in February.

The course was attended by 54 people representing the following divisions or organizations:

MIT Practice School	10
Mathematics Panel	8
Applied Nuclear Physics	10
Biology	5
K-25	4
Operations	2
AEC	1
Chemical Technology	1
Director's Division	1
Health Physics	1
Physics	1
Pilot Plant	1
REED	1
Stable Isotopes	1
Aircraft Nuclear Propulsion	4
Electronuclear Research	2
Aircraft Reactor Engineering	1

PUBLICATIONS

S. E. Atta, "Effect of Propagated Error on Inverse of Hilbert Matrix," *J. Assoc. Computing Mach.* **4**, 36–40 (1957).

R. C. F. Bartels, *Topics on the Numerical Solution of Partial Differential Equations*, ORNL-2162 (Sept. 13, 1956).

R. G. Cornell, *A New Estimation Procedure for Linear Combinations of Exponentials*, ORNL-2120 (Aug. 17, 1956).

A. C. Downing, Jr., and A. S. Householder, "Some Inverse Characteristic Value Problems," *J. Assoc. Computing Mach.* **3**, 203–207 (1956).

C. L. Gerberich, *On the Solution of a Fredholm Integral Equation*, ORNL-2241 (Dec. 19, 1956).

T. W. Hildebrandt, *I. Iterative Methods for the Approximate Solution of Linear Algebraic Systems, II. Self-Adjointness in One-Group Multi-Region Diffusion Problems*, ORNL-2146 (Aug. 24, 1956).

A. S. Householder, *Generated Error in the Solution of Certain Partial Difference Equations*, ORNL-2230 (Dec. 12, 1956).

A. S. Householder, "Solving Problems with Digital Computers," *Computers and Automation* **5**(7), 6–9, 15 (1956).

A. S. Householder, "On the Convergence of Matrix Iterations," *J. Assoc. Computing Mach.* **3**, 314–324 (1956).

A. S. Householder, "Solving Problems with a Digital Computer," *Control Engineering* 4(1), 99–105 (1957).

A. S. Householder, "Numerical Mathematics from the Viewpoint of Electronic Digital Computers," *Electronic Digital Computers and Information Processing* **4**, 21–25 (1956).

A. S. Householder, "Opportunities for Mathematicians in Government Sponsored Research," *Career Opportunities in Physics, Mathematics and Engineering,* Liberty National Life Insurance Company, Birmingham, Alabama.

A. S. Householder, "Retiring Presidential Address to the ACM, Los Angeles, August 27–29, 1956," *J. Assoc. Computing Mach.* **4**, 1–4 (1957).

S. N. Roy and M. A. Kastenbaum, "On the Hypothesis of 'No-Interaction' in a Multiway Contingency Table," *Ann. Mathematical Statistics* **27**(3), 749–757 (1956).

A. W. Kimball, "Approximate Confidence Intervals for Specific Locus Mutation Rates," *Am. Naturalist* **90**, 369–376 (1956).

G. W. Medlin, "On limits of the Real Characteristic Roots of Matrices with Real Elements," *Proc. Am. Mathematical Soc.* **7**, 912–917 (1956).